

QoS-Based Linear Transceiver Optimization for Full-Duplex Multi-User Communications

Tsung-Hui Chang, Ya-Feng Liu and Shih-Chun Lin

Abstract

In this paper, we consider a multi-user wireless system with one full duplex (FD) base station (BS) serving a set of half duplex (HD) mobile users. To cope with the in-band self-interference (SI) and co-channel interference, we formulate a quality-of-service (QoS) based linear transceiver design problem. The problem jointly optimizes the downlink (DL) and uplink (UL) beamforming vectors of the BS and the transmission powers of UL users so as to provide both the DL and UL users with guaranteed signal-to-interference-plus-noise ratio performance, using a minimum UL and DL transmission sum power. The considered system model not only takes into account noise caused by non-ideal RF circuits, analog/digital SI cancellation but also constrains the maximum signal power at the input of the analog-to-digital converter (ADC) for avoiding signal distortion due to finite ADC precision. The formulated design problem is not convex and challenging to solve in general. We first show that for a special case where the SI channel estimation errors are independent and identically distributed, the QoS-based linear transceiver design problem is globally solvable by a polynomial-time bisection algorithm. For the general case, we propose a suboptimal algorithm based on alternating optimization (AO). The AO algorithm is guaranteed to converge to a Karush-Kuhn-Tucker solution. To reduce the complexity of the AO algorithm, we further develop a fixed-point method by extending the classical uplink-downlink duality in HD systems to the FD system. Simulation results are presented to demonstrate the performance of the proposed algorithms and the comparison with HD systems.

Keywords— Beamforming, Full duplex system, Multi-user communications, Alternating optimization

EDICS: SAM-BEAM, SPC-INTF, SPC-APPL, SAM-APPL.

T.-H. Chang is the corresponding author. He is with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Shenzhen, China (e-mail: tsunghui.chang@ieee.org)

Y.-F. Liu is with the State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China (e-mail: yafliu@lsec.cc.ac.cn).

S.-C. Lin is with the Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan (email: sclin@mail.ntust.edu.tw).

The work of T.-H. Chang is supported by NSFC, China, Grant No. 61571385. The work of Y.-F. Liu is supported by the NSFC, China, Grant No. 11671419 and 11631013. The work of S.-C. Lin is supported by MOST, Taiwan, Grant No. MOST 104-2628-E-011-008-MY3.

Part of this work is presented in IEEE GLOBECOM 2016 [1].

I. INTRODUCTION

The next generation wireless communication systems target at ten times faster transmission rates and much shorter latency than the current 4G system. In addition to more powerful coding schemes and advanced multiple antenna techniques, the full duplex (FD) technique has also been considered as a solution with great potential to reach the target [2]. Ideally, a FD system can double the spectral efficiency compared to the conventional half duplex (HD) systems since it allows the node to transmit and receive signals at the same time and over the same frequency [3]. In practice, however, simultaneous transmission and reception cause severe self-interference (SI) which might greatly limit the system performance. Fortunately, recent advances in analog and digital SI cancellation (SIC) techniques have made the FD technique successfully implemented in bi-directional [4], relay [5] and WiFi [6]–[8] systems.

Recently, there have been of great interest to consider the FD techniques in the multi-user cellular systems [9]. Specifically, in such scenarios, a FD base station (BS) can serve both the downlink mobile users (DMUs) and uplink mobile users (UMUs) simultaneously. However, new challenges arise as not only the BS suffers from the SI, but also the DMUs are interfered by UMUs. This new form of uplink-to-downlink (UL-to-DL) co-channel interference (CCI) could become the performance bottleneck if not appropriately mitigated. In fact, the SI and UL-to-DL CCI couples the DL and UL transmissions, and therefore unlike the HD systems where the UL and DL transmissions can be designed separately, the FD system must consider a joint UL and DL transmission design. There have been considerable efforts studying the FD joint design problems in the literature; see, e.g., [10]–[17]. Most of the works have focused on the multiple-input multiple-output (MIMO) scenarios by assuming that both the BS and MUs are equipped with multiple antennas. For example, [11]–[14], [16] have studied beamforming/precoding and resource allocation algorithms for network sum rate or energy efficiency maximization, while [10], [17] have considered algorithms for providing the MUs with guaranteed quality-of-service (QoS).

In this paper, we consider a multiple-input single-output (MISO) scenario where each of the MUs has a single antenna. The multi-antenna FD BS employs transmit beamforming for DL transmission [18]. Unlike [11]–[14], [16], which assumed the optimal non-linear receiver for UL signal detection, we consider a linear receive beamforming scheme for signal-user detection at the BS [18]. Under these settings, we formulate a QoS-based linear transceiver design problem that minimizes the sum of UL and DL transmission powers subject to constraints that guarantee minimum signal-to-interference-plus-noise ratio (SINR) requirement of MUs. Like the MIMO formulations considered in [11]–[14], [16], the MISO QoS-based linear transceiver problem is still difficult to solve [17] due to the coupled DL and

UL transmissions. We are interested in such problem formulation for two reasons. First, the QoS based formulation is more fundamental as the solution to the problem can be used for other design formulations such as the max-min-fairness designs [19] and rate region characterization. Second, in HD systems, the DL QoS-based beamforming problem [20] and UL QoS-based receive beamforming and power control problem [21] are known polynomial-time solvable, though they are not convex problems in their original forms. Specifically, there exists an uplink-downlink duality (UDD) [19], [22], [23] between the UL and DL problems and the two problems can be efficiently solved by a fixed-point iterative method [19], [22]. Therefore, it is interesting to see whether these elegant results can be generalized to the FD systems. The main contributions of this paper are threefold:

- **Practical problem formulation:** As motivated by the circuit designs in [7], [8], [24], we consider a system model that not only takes into account the noise caused by non-ideal RF chains but also the capability of analog and digital SIC schemes. The proposed formulation based on this practical model thus captures the impact of these relevant system parameters on the network performance. In particular, different from the existing works [10]–[17], the formulated QoS-based design problem explicitly constrains the signal power level at the input of analog-to-digital converter (ADC), for avoiding signal distortion due to limited ADC precision. Such constraint is critical to the design of a FD transceiver [24], but has not been explicitly considered in the literature.
- **Polynomial-time solvable subclass:** While the QoS-based linear transceiver problem is non-convex and difficult to solve in general, we identify one intriguing case of the problem that is globally solvable. Specifically, we show that if the SI channel estimation error (i.e., the residual SI channel after analog SIC) has independent and identically distributed (i.i.d.) elements, then the non-convex problem can actually be globally solved by a bisection algorithm. The algorithm itself reveals interesting insights into the solution structure of the considered problem.
- **Efficient suboptimal algorithm:** In particular, our analysis suggests that it is unlikely to solve the QoS-based linear transceiver design problem in a convex fashion with respect to all the variables of the DL beamformer, UL beamformer and UMUs' transmission powers. In light of this, to handle the considered design problem in general, we propose an alternating optimization (AO) based suboptimal algorithm, which iteratively optimizes the UL beamformer followed by optimizing the DL beamformer and UMUs' transmission powers until convergence. Moreover, we generalize the UDD in HD systems [19], [22], [23] to the FD system and developed a new fixed-point algorithm to improve the computational efficiency of the proposed AO algorithm.

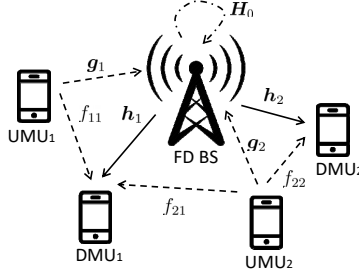


Fig. 1: A wireless network with one FD BS and multiple MUs.

Simulation results are presented to illustrate the performance of the proposed algorithms under various settings and comparison with the HD system.

Synopsis: Section II presents the FD system model and the considered QoS-based design problem. Some existing results in HD systems are also briefly reviewed. Section III considers a special case of the considered problem and proposed a globally optimal bisection algorithm. In Section IV, an AO algorithm is proposed, and the UDD and fixed-point algorithm for the FD systems are presented. Simulation results are given in Section V and conclusions are drawn in Section VI.

Notations: $\text{diag}(\mathbf{A})$ is a diagonal matrix with its diagonal elements equal to the diagonal elements of matrix \mathbf{A} ; $\text{diag}(\{a_i\}_i)$ and $\text{diag}(\mathbf{a})$, where $\mathbf{a} = [a_1, \dots, a_m]^T$, both represent a diagonal matrix with a_i 's being the diagonal elements; we also denote $|\mathbf{a}|^2 = [|a_1|^2, \dots, |a_m|^2]^T$; \mathbf{e}_n denotes the elementary vector with one in the n th entry and zero otherwise. $\mathbf{a} \succeq \mathbf{b}$ denotes element-wise inequality for vectors \mathbf{a} and \mathbf{b} , while $\mathbf{A} \succeq (\succ) \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is a positive semidefinite (positive definite) matrix. $\lambda_{\max}(\mathbf{A})$ denotes the maximum eigenvalue of matrix \mathbf{A} . $\text{vec}(\mathbf{A})$ is a vector obtained by stacking the columns of \mathbf{A} . Finally, \mathbf{I}_n denotes the n by n identity matrix and $\mathbb{E}(\cdot)$ denotes the statistical expectation operator.

II. SIGNAL MODEL AND PROBLEM FORMULATION

A. FD Signal Model

As shown in Figure 1, we consider a wireless system with one BS and a set of K DMUs and L UMUs. The DMUs want to receive information signal from the BS whereas the UMUs want to transmit information signals to the BS. We assume that a FD BS, which is equipped with N_t antennas ($N_t \geq 1$), is capable of communicating with the DMUs and UMLs at the same time and over a common spectrum. The DMUs and UMUs are assumed to be HD and have a single antenna.

We present a FD signal model which is motivated by the circuit designs proposed in [8], [24] and the signal model in [4]. In particular, the proposed model explicitly accounts for the effects of analog SIC,

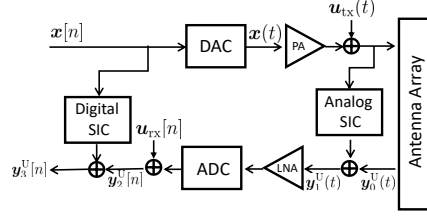


Fig. 2: A block diagram of the transmitter and receiver chains of the FD BS.

digital SIC [8], [24] as well as transmitter and receiver noises caused by non-linear circuit components [4], which therefore allows for detailed assessment of the FD system performance. The proposed model is illustrated in Figure 2. Let $\mathbf{x}[n] = \sum_{i=1}^K \mathbf{w}_i s_i^D[n]$ be the (discrete-time) signal transmitted by the BS to DMUs, where $s_i^D[n] \in \mathbb{C}$ is the (independent, zero mean and unit power) information signal for DMU i , and $\mathbf{w}_i \in \mathbb{C}^{N_t}$ is the associated beamforming vector, for all $i \in \mathcal{K} \triangleq \{1, \dots, K\}$. As shown in Figure 2, due to non-ideal transmitter RF chain (e.g., non-linearity of digital-to-analog converter (DAC) and power amplifier), the continuous-time information signal $\mathbf{x}(t) = \sum_{i=1}^K \mathbf{w}_i s_i^D(t)$ is corrupted by a transmitter noise signal $\mathbf{u}_{tx}(t) \in \mathbb{C}^{N_t}$. Following [4], we model $\mathbf{u}_{tx}(t)$ as a complex Gaussian random process with zero mean and covariance matrix $\beta_1 \text{diag}(\sum_{i=1}^K \mathbf{w}_i \mathbf{w}_i^H)$ at each time t , i.e.,

$$\mathbf{u}_{tx}(t) \sim \mathcal{CN}\left(\mathbf{0}, \beta_1 \text{diag}\left(\sum_{i=1}^K \mathbf{w}_i \mathbf{w}_i^H\right)\right), \quad (1)$$

where $\beta_1 \ll 1$ is a constant. The noise $\mathbf{u}_{tx}(t)$ is assumed to be independent of $\mathbf{x}(t)$ and the receiver noise. The combined signal $\mathbf{x}(t) + \mathbf{u}_{tx}(t)$ is then transmitted to the DMUs through the antenna array.

1) *Downlink user SINR*: Since the DMUs and UMUs communicate with the BS simultaneously, in addition to the signal from the BS, the DMUs also receive signals from the UMUs. Let $\mathbf{h}_i \in \mathbb{C}^{N_t}$ be the channel vector between the BS and the i th DMU, and let $f_{ji} \in \mathbb{C}$ be the channel coefficient from the UMU j to DMU i (see Figure 1). Moreover, denote $p_\ell^U > 0$ as the transmission power of the ℓ th UMU, and $s_\ell^U(t) \in \mathbb{C}$ as the (zero mean, unit power) UL information signal, for all $\ell \in \mathcal{L} \triangleq \{1, \dots, L\}$. Then the received signal of each DMU i is given by

$$y_i^D(t) = \mathbf{h}_i^H \left(\sum_{k=1}^K \mathbf{w}_k s_k^D(t) + \mathbf{u}_{tx}(t) \right) + \sum_{j=1}^L f_{ji} \sqrt{p_j^U} s_j^U(t) + z_i^D(t), \quad \forall i \in \mathcal{K}, \quad (2)$$

where $z_i^D(t) \in \mathbb{C}$ is the additive white Gaussian noise (AWGN) following $\mathcal{CN}(0, \sigma_i^2)$. By (2), the SINR of DMU i can be shown as

$$\text{SINR}_i^D(\mathbf{W}, \mathbf{p}^U) = \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2}{\sum_{k \neq i}^K |\mathbf{h}_i^H \mathbf{w}_k|^2 + \beta_1 \mathbf{h}_i^H \text{diag}(\sum_{k=1}^K \mathbf{w}_k \mathbf{w}_k^H) \mathbf{h}_i + \hat{\sigma}_i^2(\mathbf{p}^U)}, \quad (3)$$

for all $i \in \mathcal{K}$, where $\mathbf{W} \triangleq [\mathbf{w}_1, \dots, \mathbf{w}_K]$, $\mathbf{p}^U \triangleq [p_1^U, \dots, p_L^U]^T$ and $\hat{\sigma}_i^2(\mathbf{p}^U) \triangleq \sum_{j=1}^L p_j^U |f_{ji}|^2 + \sigma_i^2$. Note that the second and third terms in the denominator of (3) are due to the transmitter noise $\mathbf{u}_{\text{tx}}(t)$, the UL-to-DL CCI $\sum_{j=1}^L f_{ji} \sqrt{p_j^U} s_j^U(t)$ and the additive noise. Let $\gamma_i^D > 0$ be the target SINR for DMU i . Then the SINR constraint for DMU i can be written as

$$\begin{aligned} \text{SINR}_i^D(\mathbf{W}, \mathbf{p}^U) &\geq \gamma_i^D \\ \iff \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2 / \rho_i^D}{\sum_{k=1}^K \mathbf{w}_k^H \tilde{\mathbf{H}}_i \mathbf{w}_k + \hat{\sigma}_i^2(\mathbf{p}^U)} &\geq 1, \end{aligned} \quad (4)$$

where $\tilde{\mathbf{H}}_i \triangleq \mathbf{h}_i \mathbf{h}_i^H + \beta_1 \text{diag}(|\mathbf{h}_i|^2)$, for all $i \in \mathcal{K}$, and $\frac{1}{\rho_i^D} \triangleq 1 + 1/\gamma_i^D$.

2) *Uplink user SINR*: Denote $\mathbf{H}_0 \in \mathbb{C}^{N_t \times N_t}$ as the SI channel matrix and $\mathbf{g}_\ell \in \mathbb{C}^{N_t}$ as the channel vector from the ℓ th UMU to the BS (see Figure 1). The signal received by the BS thus can be expressed as

$$\mathbf{y}_0^U(t) = \sum_{\ell=1}^L \mathbf{g}_\ell \sqrt{p_\ell^U} s_\ell^U(t) + \mathbf{H}_0(\mathbf{x}(t) + \mathbf{u}_{\text{tx}}(t)) + \mathbf{z}^U(t), \quad (5)$$

where the second term in the right hand side (RHS) is the SI and the third term $\mathbf{z}^U(t)$ is the AWGN following $\mathcal{CN}(\mathbf{0}, \sigma_z^2 \mathbf{I}_{N_t})$. The SI power is in general much stronger than the signals transmitted from the UMUs and the AWGN, and therefore the SI has to be suppressed in order to decode the desired information data properly. However, simply mitigating the SI in the digital domain is insufficient. In fact, the SI power could be so large such that the receiver RF chain gets saturated as the dynamics of $\mathbf{y}_0^U(t)$ may be out of the range that the ADC can support [24]. In view of this, SIC has to be carried out in the analog domain before ADC [24], as shown in Figure 2. Let $\mathbf{H}_0 = \hat{\mathbf{H}}_0 + \Phi_0$, where $\hat{\mathbf{H}}_0$ and Φ_0 respectively represent the channel estimate of \mathbf{H}_0 and the associated estimation error matrix. Suppose that the analog SIC subtracts the estimated SI signal $\hat{\mathbf{H}}_0(\mathbf{x}(t) + \mathbf{u}_{\text{tx}}(t))$ from $\mathbf{y}_0^U(t)$. Then the signal before ADC is given by

$$\mathbf{y}_1^U(t) = \sum_{\ell=1}^L \mathbf{g}_\ell \sqrt{p_\ell^U} s_\ell^U(t) + \Phi_0(\mathbf{x}(t) + \mathbf{u}_{\text{tx}}(t)) + \mathbf{z}^U(t). \quad (6)$$

After ADC, we obtain the following discrete-time signal

$$\begin{aligned} \mathbf{y}_2^U[n] &= \sum_{\ell=1}^L \mathbf{g}_\ell \sqrt{p_\ell^U} s_\ell^U[n] + \Phi_0(\mathbf{x}[n] + \mathbf{u}_{\text{tx}}[n]) \\ &\quad + \mathbf{u}_{\text{rx}}[n] + \mathbf{z}^U[n]. \end{aligned} \quad (7)$$

Here, $\mathbf{u}_{\text{rx}}[n]$ is the noise caused by the non-ideal receiver RF chain and follows

$$\mathbf{u}_{\text{rx}}[n] \sim \mathcal{CN}(\mathbf{0}, \beta_2 \text{diag}(\mathbb{E}[\mathbf{y}_1^U[n](\mathbf{y}_1^U[n])^H])), \quad (8)$$

where $\beta_2 \ll 1$ [4]. Before data detection, digital SIC is further carried out for $\mathbf{y}_2^U[n]$. According to [8], [24], it is possible to suppress the linear SI components $\Phi_0 \mathbf{x}[n]$ and non-linear components $\Phi_0 \mathbf{u}_{\text{tx}}[n] + \mathbf{u}_{\text{rx}}[n]$ separately. For simplicity, we model that the linear SI power can be reduced by a factor of $\delta_1 \ll 1$ and the non-linear SI power can be reduced by a factor of $\delta_2 \ll 1$, due to the digital SIC. Therefore, the signal at the output of the digital SIC is given by

$$\begin{aligned} \mathbf{y}_3^U[n] &= \sum_{\ell=1}^L \mathbf{g}_\ell \sqrt{p_\ell^U} s_\ell^U[n] + \sqrt{\delta_1} \Phi_0 \mathbf{x}[n] \\ &\quad + \sqrt{\delta_2} (\Phi_0 \mathbf{u}_{\text{tx}}[n] + \mathbf{u}_{\text{rx}}[n]) + \mathbf{z}^U[n]. \end{aligned} \quad (9)$$

To detect $s_\ell^U[n]$, the BS applies a linear receive beamformer, denoted by $\mathbf{v}_\ell \in \mathbb{C}^{MN_R}$, to $\mathbf{y}_3^U[n]$, for all $\ell \in \mathcal{L}$. The SINR at the output of beamformer for UMU ℓ is thus given by

$$\begin{aligned} \text{SINR}_\ell^U(\mathbf{W}, \mathbf{p}^U, \mathbf{v}_\ell) &= \\ &= \frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2}{\sum_{j \neq \ell}^L p_j^U |\mathbf{v}_\ell^H \mathbf{g}_j|^2 + \mathbf{v}_\ell^H \mathbb{E}[\Sigma(\mathbf{W}, \Phi_0)] \mathbf{v}_\ell + \sigma_z^2 \|\mathbf{v}_\ell\|_2^2}, \end{aligned} \quad (10)$$

where

$$\begin{aligned} \mathbf{v}_\ell^H \mathbb{E}[\Sigma(\mathbf{W}, \Phi_0)] \mathbf{v}_\ell &\triangleq \delta_1 \mathbf{v}_\ell^H \mathbb{E}[\Phi_0 \mathbf{x}(t) \mathbf{x}^H(t) \Phi_0^H] \mathbf{v}_\ell \\ &\quad + \delta_2 \mathbf{v}_\ell^H \mathbb{E}[\Phi_0 \mathbf{u}_{\text{tx}}[n] \mathbf{u}_{\text{tx}}^H[n] \Phi_0^H] \mathbf{v}_\ell + \delta_2 \mathbf{v}_\ell^H \mathbb{E}[\mathbf{u}_{\text{rx}}[n] \mathbf{u}_{\text{rx}}^H[n]] \mathbf{v}_\ell. \end{aligned} \quad (11)$$

It is shown in Appendix A that (11) can be compactly expressed as

$$\mathbf{v}_\ell^H \mathbb{E}[\Sigma(\mathbf{W}, \Phi_0)] \mathbf{v}_\ell = (\delta_2 \beta_2) \mathbf{v}_\ell^H \left(\sum_{j=1}^L p_j^U \text{diag}(|\mathbf{g}_j|^2) \right) \mathbf{v}_\ell + \mathbf{v}_\ell^H \mathbf{\Omega}(\mathbf{W}, \mathbf{R}_{\Phi_0}) \mathbf{v}_\ell, \quad (12)$$

$$= (\delta_2 \beta_2) \mathbf{v}_\ell^H \left(\sum_{j=1}^L p_j^U \text{diag}(|\mathbf{g}_j|^2) \right) \mathbf{v}_\ell + \sum_{k=1}^K \mathbf{w}_k^H \mathbf{\Lambda}(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) \mathbf{w}_k, \quad (13)$$

using the matrices $\mathbf{\Omega}(\mathbf{W}, \mathbf{R}_{\Phi_0})$ and $\mathbf{\Lambda}(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0})$ defined in (A.7) and (A.8), where $\mathbf{R}_{\Phi_0} \triangleq \mathbb{E}[\text{vec}(\mathbf{\Phi}_0)\text{vec}^H(\mathbf{\Phi}_0)]$ is the correlation matrix of vectorized $\mathbf{\Phi}_0$. Let us define

$$\tilde{\mathbf{G}}_j \triangleq \mathbf{g}_j \mathbf{g}_j^H + \delta_2 \beta_2 \text{diag}(|\mathbf{g}_j|^2), \quad (14)$$

$$\tilde{\sigma}_z^2 \triangleq (1 + \delta_2 \beta_2) \sigma_z^2, \quad (15)$$

and let $\gamma_\ell^U > 0$ be the SINR target of UMU ℓ for all $\ell \in \mathcal{L}$. Then either using (12) or (13), the SINR constraint $\text{SINR}_\ell^U \geq \gamma_\ell^U$ for UMU ℓ can be written as

$$\begin{aligned} \text{SINR}_\ell^U(\mathbf{W}, \mathbf{p}^U, \mathbf{v}_\ell) &\geq \gamma_\ell^U \\ \Leftrightarrow \frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \mathbf{v}_\ell^H \mathbf{\Omega}(\mathbf{W}, \mathbf{R}_{\Phi_0}) \mathbf{v}_\ell + \tilde{\sigma}_z^2 \|\mathbf{v}_\ell\|_2^2} &\geq 1, \end{aligned} \quad (16)$$

$$\Leftrightarrow \frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \sum_{k=1}^K \mathbf{w}_k^H \mathbf{\Lambda}(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) \mathbf{w}_k + \tilde{\sigma}_z^2 \|\mathbf{v}_\ell\|_2^2} \geq 1, \quad (17)$$

where $\frac{1}{\rho_\ell^U} \triangleq 1 + 1/\gamma_\ell^U$. It will be seen shortly that both formulations in (16) and (17) are useful in the development of the proposed algorithms.

3) *Constraint on ADC Input Signal Power:* As mentioned, the signal dynamics should not be out of the range that the ADC can support as, otherwise, signal distortion can occur. Therefore, it is required to constrain the signal power of $\mathbf{y}_1^U(t)$ at the input of the ADC. Let $\gamma^{\text{ADC}} > 0$ denote the maximum tolerable ADC input signal power. The ADC power constraint for the receiver RF chains are given by

$$\text{ADC}(\mathbf{W}, \mathbf{p}^U) \triangleq \text{diag}(\mathbb{E}\{\mathbf{y}_1^U(t)(\mathbf{y}_1^U(t))^H\}) \preceq \gamma^{\text{ADC}} \mathbf{I}_{N_t}. \quad (18)$$

It can be shown that, for $n = 1, \dots, N_t$,

$$\begin{aligned} [\text{ADC}(\mathbf{W}, \mathbf{p}^U)]_{n,n} &= \mathbf{e}_n^T \mathbb{E}\{\mathbf{\Phi}_0 \mathbf{W} \mathbf{W}^H \mathbf{\Phi}_0^H\} \mathbf{e}_n \\ &+ \beta_1 \mathbf{e}_n^T \mathbb{E}\{\mathbf{\Phi}_0 \text{diag}(\mathbf{W} \mathbf{W}^H) \mathbf{\Phi}_0^H\} \mathbf{e}_n + \sum_{j=1}^L p_j^U |\mathbf{e}_n^T \mathbf{g}_j|^2 + \sigma_z^2, \\ &= \sum_{k=1}^K \mathbf{w}_k^H \mathbf{\Upsilon}_n(\{\mathbf{R}_{\Phi_0, m}\}) \mathbf{w}_k + \sum_{j=1}^L p_j^U |\mathbf{e}_n^T \mathbf{g}_j|^2 + \sigma_z^2, \end{aligned} \quad (19)$$

where $\mathbf{\Upsilon}_n(\{\mathbf{R}_{\Phi_0, m}\}) \triangleq \bar{\mathbf{R}}_{\Phi_0, n} + \beta_1 \text{diag}(\{\mathbf{e}_n^T \mathbf{R}_{\Phi_0, m} \mathbf{e}_n\}_m)$, and $\bar{\mathbf{R}}_{\Phi_0, n}$ and $\mathbf{R}_{\Phi_0, m}$ are defined in (A.9).

B. Proposed Problem Formulation

Our goal is to design the UL transmission powers $\{p_\ell^U\}$ and UL and DL beamformers $\{\mathbf{v}_\ell\}$, $\{\mathbf{w}_k\}$ so that the transmission power of the network (including the BS and the UMs) is minimized subject to user

SINR constraints and the ADC input power constraint. Mathematically, the QoS-based linear transceiver design problem is formulated as follows

$$(P) \quad \min_{\substack{\{\mathbf{w}_k\}, \{\mathbf{v}_\ell\}, \\ \{p_\ell^U \geq 0\}}} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \sum_{\ell=1}^L p_\ell^U \quad (20a)$$

$$\text{s.t.} \quad \text{SINR}_i^D(\mathbf{W}, \mathbf{p}^U) \geq \gamma_i^D, \quad i \in \mathcal{K}, \quad (20b)$$

$$\text{SINR}_\ell^U(\mathbf{W}, \mathbf{p}^U, \mathbf{v}_\ell) \geq \gamma_\ell^U, \quad \ell \in \mathcal{L}, \quad (20c)$$

$$\text{ADC}(\mathbf{W}, \mathbf{p}^U) \preceq \gamma^{\text{ADC}} \mathbf{I}_{N_t}, \quad (20d)$$

$$\|\mathbf{v}_\ell\|_2 = 1, \quad \ell \in \mathcal{L}. \quad (20e)$$

Unfortunately, problem (P) is non-convex and difficult to solve. Specifically, the UL SINR constraints and DL SINR constraints are coupled with each other due to the FD BS, which makes (P) drastically different from the traditional design problems in HD systems [18]. In subsequent sections, we present two methods to handle (P). Firstly, we show that for a special case of (P), the problem is globally solvable in a polynomial-time complexity. Secondly, for general (P), we propose an efficient AO method to solve it approximately.

C. Review of Half-Duplex BF Solutions

Before studying the methods for solving the FD design problem (P), let us review some existing results about the QoS-based design problems in a HD system. These results will be used in the development of the proposed methods for solving (P).

In the HD system, the BS serves the DMUs and UMUs separately, either in different time slots or over distinct frequency bands. Thus, there is no SI (i.e., the term $\sum_{k=1}^K \mathbf{w}_k^H \mathbf{\Lambda}(\mathbf{v}_\ell, \mathbf{\Phi}_0) \mathbf{w}_k$ in (17)) and no UL-to-DL CCI (i.e., the term $\sum_{j=1}^L p_j^U |f_{ji}|^2$ in (4)). The ADC input power constraint is also not necessary as the signal power at the input of ADC is generally small in the absence of SI. Therefore, the DL design problem (HD-DL)

$$\min_{\{\mathbf{w}_k\}} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \quad (21a)$$

$$\text{s.t.} \quad \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2 / \rho_i^D}{\sum_{k=1}^K \mathbf{w}_k^H \tilde{\mathbf{H}}_i \mathbf{w}_k + \sigma_i^2} \geq 1, \quad k \in \mathcal{K}, \quad (21b)$$

and UL design problem (HD-UL)

$$\min_{\{\mathbf{v}_\ell\}, \{p_\ell^U \geq 0\}} \sum_{\ell=1}^L p_\ell^U \quad (22a)$$

$$\text{s.t.} \quad \frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \mathbf{g}_j \mathbf{g}_j^H \mathbf{v}_\ell + \sigma_z^2 \|\mathbf{v}_\ell\|_2^2} \geq 1, \quad (22b)$$

$$\|\mathbf{v}_\ell\|_2 = 1, \quad \ell \in \mathcal{L}, \quad (22c)$$

can respectively be deduced from (P), and optimized independently [18].

While both (HD-DL) and (HD-UL) appear non-convex problems, it is well known that both of them own certain hidden convexity and can be globally solved in a polynomial-time complexity. Specifically, (HD-UL) is shown equivalent to a convex semidefinite program (SDP).

Lemma 1 [21] *Suppose that (HD-UL) is feasible. Then (HD-UL) is equivalent to the following SDP*

$$\begin{aligned} \max_{\{p_\ell^U \geq 0\}} \quad & \sum_{\ell=1}^L p_\ell^U \\ \text{s.t.} \quad & \sum_{j=1}^L p_j^U \mathbf{g}_j \mathbf{g}_j^H + \sigma_z^2 \mathbf{I}_{N_t} \succeq \left(\frac{p_\ell^U}{\rho_\ell^U} \right) \mathbf{g}_\ell \mathbf{g}_\ell^H, \quad \forall \ell \in \mathcal{L}, \end{aligned} \quad (23)$$

and therefore (HD-UL) is polynomial-time solvable.

Lemma 2 [21, Proposition 3.1] [25] *Suppose that (HD-UL) is feasible. The optimal power solution of problem (HD-UL) is unique and satisfies*

$$\left(\frac{p_\ell^U}{\rho_\ell^U} \right) \mathbf{g}_\ell^H \left(\sum_{j=1}^L p_j^U \mathbf{g}_j \mathbf{g}_j^H + \sigma_z^2 \mathbf{I}_{N_t} \right)^{-1} \mathbf{g}_\ell = 1, \quad \forall \ell \in \mathcal{L}. \quad (24)$$

The optimal beamformers are given by

$$\mathbf{v}_\ell = \frac{\mathbf{v}_\ell}{\|\mathbf{v}_\ell\|}, \quad \tilde{\mathbf{v}}_\ell = \left(\sum_{j=1}^L p_j^U \mathbf{g}_j \mathbf{g}_j^H + \sigma_z^2 \mathbf{I}_{N_t} \right)^{-1} \mathbf{g}_\ell, \quad \forall \ell \in \mathcal{L}. \quad (25)$$

In fact, it can be verified that at the optimum, constraint (22b) must hold with equality, which leads to (24). Thus, Lemma 2 says that the L nonlinear system of equations of (24) has a unique solution, uniquely determined by \mathbf{g}_j 's, \mathbf{g}_ℓ 's, γ_ℓ^U 's and σ_z^2 . Moreover, the optimal $\{p_\ell^U\}$ can be obtained by solving equations (24) using a fixed-point method [19], [22].

The DL problem (HD-DL) can be solved by considering an equivalent second-order cone program (SOCP) [19] or a SDP [20] which is obtained by a semidefinite relaxation (SDR) technique [26]. Both SOCP and SDP are convex problems and are efficiently solvable by off-the-shelf solvers. The fixed-point

method for UL problems can also be used to solve the DL problem (HD-DL), through a powerful UDD [18], [19], [27].

Lemma 3 [18] *Suppose that (HD-DL) is feasible. (HD-DL) has a virtual UL counterpart as follows*

$$\min_{\{\tilde{\mathbf{w}}_i\}, \{\lambda_i \geq 0\}} \sum_{i=1}^K \lambda_i \sigma_i^2 \quad (26a)$$

$$\text{s.t. } \frac{\lambda_i |\tilde{\mathbf{w}}_i^H \mathbf{h}_i|^2 / \rho_i^D}{\sum_{k=1}^K \lambda_k \tilde{\mathbf{w}}_i^H \tilde{\mathbf{H}}_k \tilde{\mathbf{w}}_i + \|\tilde{\mathbf{w}}_i\|_2^2} \geq 1, \quad (26b)$$

$$\|\tilde{\mathbf{w}}_i\|_2 = 1, \quad i \in \mathcal{K}, \quad (26c)$$

which has the same optimal objective value as (HD-DL).

In Section IV-B, we will show that such UDD can be generalized to the FD system.

III. OPTIMAL SOLUTION FOR (P) WITH I.I.D. SI CHANNEL ESTIMATION ERROR

In this section, we consider a special case of (P) by assuming that elements of the SI channel estimation error are i.i.d. Specifically, \mathbf{R}_{Φ_0} is assumed to be $\mathbf{R}_{\Phi_0} = \sigma_{\Phi_0}^2 \mathbf{I}_{N_t^2}$ for some $\sigma_{\Phi_0}^2 > 0$. Under this assumption, one can show that $\Omega(\mathbf{W}, \mathbf{R}_{\Phi_0})$ in the denominator of (16) reduces to $\xi \sum_{i=1}^K \|\mathbf{w}_i\|_2^2 \mathbf{I}_{N_t}$, where $\xi \triangleq (\delta_1 + \delta_2 \beta_1 + \delta_2 \beta_2 (1 + \beta_1)) \sigma_{\Phi_0}^2$, and $\Upsilon_n(\{\mathbf{R}_{\Phi_0, m}\})$ in the ADC input power constraint (19) reduces to $\Upsilon_n(\{\mathbf{R}_{\Phi_0, m}\}) = \sigma_{\Phi_0}^2 (1 + \beta_1) \sum_{i=1}^K \|\mathbf{w}_i\|_2^2 \mathbf{I}_{N_t}$. As a result, problem (P) simplifies to

$$(P1) \quad \min_{\substack{\{\mathbf{w}_k\}, \{\mathbf{v}_\ell\}, \\ \{p_\ell^U \geq 0\}}} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \sum_{\ell=1}^L p_\ell^U \quad (27a)$$

$$\text{s.t. } \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2 / \rho_i^D}{\sum_{k=1}^K \mathbf{w}_k^H \tilde{\mathbf{H}}_i \mathbf{w}_k + \hat{\sigma}_i^2 (p^U)} \geq 1, \quad i \in \mathcal{K}, \quad (27b)$$

$$\frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \xi \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \tilde{\sigma}_z^2} \geq 1, \quad \ell \in \mathcal{L}, \quad (27c)$$

$$\overline{\text{ADC}}(\mathbf{W}, \mathbf{p}^U) \preceq \gamma^{\text{ADC}} \mathbf{I}_{N_t}, \quad (27d)$$

$$\|\mathbf{v}_\ell\|_2 = 1, \quad \ell \in \mathcal{L}, \quad (27e)$$

where $[\overline{\text{ADC}}(\mathbf{W}, \mathbf{p}^U)]_{n,n} \triangleq \sigma_{\Phi_0}^2 (1 + \beta_1) \sum_{i=1}^K \|\mathbf{w}_i\|_2^2 + \sum_{j=1}^L p_j^U |\mathbf{e}_n^T \mathbf{g}_j|^2 + \sigma_z^2$, for all $n = 1, \dots, N_t$. Note that even when $\mathbf{R}_{\Phi_0} \neq \sigma_{\Phi_0}^2 \mathbf{I}_{N_t^2}$, one can still obtain a similar formulation as (P1) by considering a worst-case SI. In particular, one can show that $\Omega(\mathbf{W}, \mathbf{R}_{\Phi_0})$ is upper bounded as $\Omega(\mathbf{W}, \mathbf{R}_{\Phi_0}) \preceq (\delta_1 + \delta_2 \beta_1 + \delta_2 \beta_2 (1 + \beta_1)) \lambda_{\max}(\mathbf{R}_{\Phi_0}) \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \mathbf{I}_{N_t}$. Besides, $\Upsilon_n(\{\mathbf{R}_{\Phi_0, m}\})$ is upper bounded as $\Upsilon_n(\{\mathbf{R}_{\Phi_0, m}\}) \preceq \lambda_{\max}(\mathbf{R}_{\Phi_0}) (1 + \beta_1) \sum_{i=1}^K \|\mathbf{w}_i\|_2^2 \mathbf{I}_{N_t}$. Therefore, by replacing $\Omega(\mathbf{W}, \mathbf{R}_{\Phi_0})$ and $\Upsilon_n(\{\mathbf{R}_{\Phi_0, m}\})$ by their

respective upper bounds, one can arrive at a similar formulation as (P1). Interestingly, as we will show shortly, despite that (P1) is still a non-convex problem, it is actually polynomial-time solvable.

A. A Globally Solvable Non-Convex Subproblem

One of the keys to the developed algorithm for solving (P1) is to consider the following problem

$$(P_\eta) \quad F(\eta) = \min_{\substack{\{\mathbf{w}_k\}, \{\mathbf{v}_\ell\}, \\ \{p_\ell^U \geq 0\}}} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \sum_{\ell=1}^L p_\ell^U \quad (28a)$$

$$\text{s.t.} \quad \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2 / \rho_i^D}{\sum_{k=1}^K \mathbf{w}_k^H \tilde{\mathbf{H}}_i \mathbf{w}_k + \hat{\sigma}_i^2(\mathbf{p}^U)} \geq 1, \quad i \in \mathcal{K}, \quad (28b)$$

$$\frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \xi \eta + \tilde{\sigma}_z^2} \geq 1, \quad \ell \in \mathcal{L}, \quad (28c)$$

$$\overline{\text{ADC}}(\mathbf{W}, \mathbf{p}^U) \preceq \gamma^{\text{ADC}} \mathbf{I}_{N_t}, \quad (28d)$$

$$\|\mathbf{v}_\ell\|_2 = 1, \quad \ell \in \mathcal{L}, \quad (28e)$$

where $F(\eta)$ denotes the optimal objective value. The difference between (P_η) and (P1) is that the term $\sum_{k=1}^K \|\mathbf{w}_k\|_2^2$ in (27c) is replaced by a parameter $\eta > 0$. Intriguingly, problem (P_η) , though not being convex, is polynomial-time solvable.

Proposition 1 *Suppose that problem (P_η) is feasible. Then, the optimal $\{\mathbf{v}_\ell, p_\ell^U\}$ of (P_η) , denoted by $\{\mathbf{v}_\ell(\eta), p_\ell^U(\eta)\}$, is the solution to the following UL problem*

$$\{\mathbf{v}_\ell(\eta), p_\ell^U(\eta)\}_{\ell=1}^L = \arg \min_{\substack{\{\mathbf{v}_\ell\}, \{p_\ell^U \geq 0\}}} \sum_{\ell=1}^L p_\ell^U \quad (29a)$$

$$\text{s.t.} \quad \frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + (\xi \eta + \tilde{\sigma}_z^2)} \geq 1, \quad \ell \in \mathcal{L}, \quad (29b)$$

$$\|\mathbf{v}_\ell\|_2 = 1, \quad \ell \in \mathcal{L}. \quad (29c)$$

and the optimal $\{\mathbf{w}_k\}$ of (P_η) , denoted by $\{\mathbf{w}_k(\eta)\}$, is the solution to the following DL problem

$$\{\mathbf{w}_k(\eta)\}_{k=1}^K = \arg \min_{\{\mathbf{w}_k\}} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \quad (30a)$$

$$\text{s.t.} \quad \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2 / \rho_i^D}{\sum_{k=1}^K \mathbf{w}_k^H \tilde{\mathbf{H}}_i \mathbf{w}_k + \hat{\sigma}_i^2(\mathbf{p}^U(\eta))} \geq 1, \quad i \in \mathcal{K}, \quad (30b)$$

$$\overline{\text{ADC}}(\mathbf{W}, \mathbf{p}^U(\eta)) \preceq \gamma^{\text{ADC}} \mathbf{I}_{N_t}. \quad (30c)$$

Both problems in (29) and (30) are polynomial-time solvable.

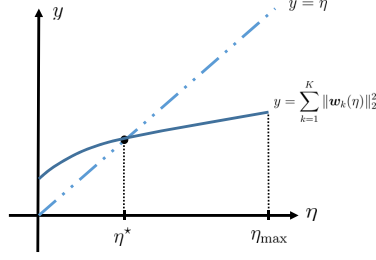


Fig. 3: Illustration of Proposition 3. It is assumed that (P1) is feasible.

Proof: Note that the UL SINR constraints (28c) must hold with equality at the optimum; otherwise, one can further reduce p_ℓ^U 's without violating the other constraints. Since $\{\mathbf{v}_\ell\}$ appears only in (28c), $\{\mathbf{v}_\ell(\eta)\}$ is the maximum SINR solution given by

$$\begin{aligned} \mathbf{v}_\ell(\eta) &= \frac{\mathbf{v}_\ell(\eta)}{\|\mathbf{v}_\ell(\eta)\|_2}, \\ \tilde{\mathbf{v}}_\ell(\eta) &= \left(\sum_{j=1}^L p_j^U(\eta) \tilde{\mathbf{G}}_j + (\xi\eta + \tilde{\sigma}_z^2) \mathbf{I}_{N_t} \right)^{-1} \mathbf{g}_\ell, \quad \forall \ell \in \mathcal{L}. \end{aligned} \quad (31)$$

So, $\{p_\ell^U(\eta)\}$ satisfies, $\forall \ell \in \mathcal{L}$,

$$\frac{p_\ell^U(\eta)}{\rho_\ell^U} \mathbf{g}_\ell^H \left(\sum_{j=1}^L p_j^U(\eta) \tilde{\mathbf{G}}_j + (\xi\eta + \tilde{\sigma}_z^2) \mathbf{I}_{N_t} \right)^{-1} \mathbf{g}_\ell = 1. \quad (32)$$

By applying Lemma 2, we see that $\{p_\ell^U(\eta)\}$ is uniquely determined by (32), regardless of the constraints in (28b) and (28d). As a result, $\{\mathbf{v}_\ell(\eta), p_\ell^U(\eta)\}$ essentially can be obtained by solving (29), which is similar to the HD UL problem (HD-UL) and is polynomial-time solvable. Once $\{\mathbf{v}_\ell(\eta), p_\ell^U(\eta)\}$ are given, it is clear that $\{\mathbf{w}_k(\eta)\}$ can be obtained by solving (30). Problem (30) is polynomial-time solvable as, similar to (HD-DL), (30) can be formulated as an SOCP [19] or solved by SDR [20]. ■

Proposition 1 is constructive as it provides a two-stage approach to solving (P $_\eta$). In the next subsection, we analyze the relation between (P $_\eta$) and (P1). Based on these results, we develop a bisection strategy to solve (P1) to global optimality.

B. Proposed Bisection Algorithm for Solving (P1)

We first have the following proposition.

Proposition 2 Suppose that (P1) is feasible. Let $\{\mathbf{w}_k^*\}$ be the optimal DL beamformers of (P1) and let

$$\eta^* \triangleq \sum_{k=1}^K \|\mathbf{w}_k^*\|_2^2. \text{ Then}$$

(a) η^* is unique;

- (b) When $\eta = \eta^*$, $\{\mathbf{w}_k(\eta)\}$ of (\mathbf{P}_η) is optimal to $(\mathbf{P1})$ and satisfies $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 = \eta^*$;
(c) $F(\eta)$ of (\mathbf{P}_η) is monotonically increasing with η .

Proof: See Appendix B. ■

Proposition 2 suggests that one can solve $(\mathbf{P1})$ by searching the unique optimal η^* of (\mathbf{P}_η) . The following result provides important insight into how η can be searched efficiently.

Proposition 3 Suppose that $(\mathbf{P1})$ is feasible and that (\mathbf{P}_η) is feasible for $\eta \in [0, \eta_{\max}]$, where $\eta_{\max} > 0$.

Then

- (a) $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ of (\mathbf{P}_η) is concave and increasing with respect to $\eta \in [0, \eta_{\max}]$.
(b) $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 > \eta$ if and only if $\eta < \eta^*$.

Moreover, $(\mathbf{P1})$ is infeasible if and only if (\mathbf{P}_η) has $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 > \eta$ for all $\eta \in [0, \eta_{\max}]$.

Proof: See Appendix C. ■

From Proposition 3, we can visualize the function $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ as in Figure 3, provided that $(\mathbf{P1})$ is feasible. Therefore, one can search the optimal η^* in a bisection fashion, by comparing the values of $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ and η . Specifically, if $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 > \eta$, η should be increased, whereas if $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 \leq \eta$, then η should be decreased. Moreover, if it happens that $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 > \eta$ for all $\eta \in [0, P_{\max}]$, where P_{\max} denotes some bisection upper bound, then it will make η converges to the bisection upper bound and thus one can declare that $(\mathbf{P1})$ is infeasible for $\sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \leq P_{\max}$. Based on these observations, we develop a bisection method for solving $(\mathbf{P1})$ in Algorithm 1. In Algorithm 1, the parameter P_{\max} denotes the bisection upper bound and is the maximum DL transmission power budget.

Algorithm 1 Proposed bisection algorithm for solving $(\mathbf{P1})$.

- 1: **Set** $\eta^{\ell b} = 0$ and $\eta^{ub} = P_{\max}$.
 - 2: **repeat**
 - 3: Set $\eta^m \leftarrow \frac{\eta^{ub} + \eta^{\ell b}}{2}$.
 - 4: Solve (\mathbf{P}_η) by solving (29) and (30) in order.
 - 5: If (\mathbf{P}_η) is infeasible or if (\mathbf{P}_η) is feasible and $\sum_{k=1}^K \|\mathbf{w}_k(\eta^m)\|_2^2 \leq \eta^m$, set $\eta^{ub} \leftarrow \eta^m$; otherwise set $\eta^{\ell b} \leftarrow \eta^m$.
 - 6: **until** $|\eta^{ub} - \eta^{\ell b}| \leq \epsilon$, where ϵ is a preset value. If $|\eta^{ub} - P_{\max}| \leq \epsilon$, then declare that $(\mathbf{P1})$ is infeasible.
-

Since both the bisection algorithm and the subproblems (29) and (30) have polynomial-time complexity, (P1) is polynomial-time solvable.

Theorem 1 *Suppose that (P1) is feasible. Then Algorithm 1 globally solves (P1) in polynomial time.*

From the proof of Proposition 3(a), it is worth noting that not only $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ but also $\sum_{\ell=1}^L p_\ell^U(\eta)$ are concave functions of η . When applying this insight to the original problem (P), it implies that the objective function $\sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \sum_{\ell=1}^L p_\ell^U$ is concave function of the SI power. Therefore, for such minimization problem, it is unlikely to solve (P) jointly with respect to all the variables $\{\mathbf{w}_k, \mathbf{v}_\ell, p_\ell^U\}$ in a convex manner. In light of this, we resort to approximation methods in the next section.

IV. KKT SOLUTIONS TO (P) VIA ALTERNATING OPTIMIZATION

In this section, we consider the general design problem (P) in (20) and propose a suboptimal method to handle (P) by alternating optimization. In the first subsection, we present the proposed AO algorithm and show that the algorithm is guaranteed to converge to a Karush-Kuhn-Tucker (KKT) point of (P). In the second subsection, we generalize the UDD in HD systems [25], [27] to the FD system and use it to develop a computationally efficient fixed-point based AO algorithm.

A. Proposed AO Algorithm

The proposed AO algorithm shares a similar strategy as the iterative algorithms proposed in [28]–[30] for joint transmit and receive beamforming optimization in HD MIMO systems. For the considered problem (P), we observe that, when the UL beamformer $\{\mathbf{v}_\ell\}$ are fixed, (P) can be recast as a convex SOCP, through simple change of variables. Moreover, when the DL beamformer $\{\mathbf{w}_k\}$ and UL power $\{p_\ell^U\}$ are fixed, $\{\mathbf{v}_\ell\}$ has a simple closed-form expression.

Specifically, let us assume that the UL beamforming vectors $\{\mathbf{v}_\ell\}$ are fixed in (P). We have

$$\min_{\{\mathbf{w}_k\}, \{p_\ell^U \geq 0\}} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \sum_{\ell=1}^L p_\ell^U \quad (33a)$$

$$\text{s.t.} \quad \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2 / \rho_i^D}{\sum_{k=1}^K \mathbf{w}_k^H \tilde{\mathbf{H}}_i \mathbf{w}_k + \hat{\sigma}_i^2(\mathbf{p}^U)} \geq 1, \quad \forall i \in \mathcal{K}, \quad (33b)$$

$$\frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \sum_{k=1}^K \mathbf{w}_k^H \mathbf{\Lambda}(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) \mathbf{w}_k + \tilde{\sigma}_z^2} \geq 1, \quad \forall \ell \in \mathcal{L}, \quad (33c)$$

$$\text{ADC}(\mathbf{W}, \mathbf{p}^U) \preceq \gamma^{\text{ADC}} \mathbf{I}_{N_t}. \quad (33d)$$

Note from (19) that (33d) are convex constraints. Besides, if $\{\mathbf{w}_k^*\}$ is an optimal solution to (33), then any phase rotated version of $\{\mathbf{w}_k^*\}$ is still an optimal solution. Let us consider in (33) the change of variables $q_\ell = \sqrt{p_\ell^U}$, $\ell \in \mathcal{L}$, and apply proper phase rotation to $\{\mathbf{w}_k\}$ so that $\mathbf{h}_k^H \mathbf{w}_k$ is real-valued for all $k \in \mathcal{K}$. Then one can equivalently write (33) as follows

$$\min_{\{\mathbf{w}_k\}, \{q_\ell \geq 0\}} \sum_{k=1}^K \|\mathbf{w}_k\|^2 + \sum_{\ell=1}^L q_\ell^2 \quad (34a)$$

$$\text{s.t. } \frac{\mathbf{h}_i^H \mathbf{w}_i}{\sqrt{\rho_i^D}} \geq \sqrt{\sum_{k=1}^K \|\tilde{\mathbf{H}}_i^{\frac{1}{2}} \mathbf{w}_k\|_2^2 + \sum_{j=1}^L q_j^2 |f_{ji}|^2 + \sigma_i^2}, \quad k \in \mathcal{K}, \quad (34b)$$

$$\frac{q_\ell |\mathbf{v}_\ell^H \mathbf{g}_\ell|}{\sqrt{\rho_\ell^U}} \geq \sqrt{\sum_{j=1}^L q_j^2 \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \sum_{k=1}^K \|\Lambda^{\frac{1}{2}}(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) \mathbf{w}_k\|_2^2 + \tilde{\sigma}_z^2}, \quad \forall \ell \in \mathcal{L}, \quad (34c)$$

$$\text{ADC}(\mathbf{W}, \mathbf{p}^U) \preceq \gamma^{\text{ADC}} \mathbf{I}_{N_t}. \quad (34d)$$

Problem (34) is an SOCP which can be solved by standard convex solvers.

On the other hand, suppose that $\{\mathbf{w}_k\}$ and $\{p_\ell^U\}$ s are fixed in (P). Then the optimal \mathbf{v}_ℓ that maximizes SINR_ℓ^U in (20c) (i.e., (16)) is the maximum SINR beamformer

$$\mathbf{v}_\ell = \frac{\mathbf{M}_\ell^{-1} \mathbf{g}_\ell}{\|\mathbf{M}_\ell^{-1} \mathbf{g}_\ell\|_2}, \quad (35)$$

where $\mathbf{M}_\ell \triangleq \sum_{j=1}^L p_j^U \tilde{\mathbf{G}}_j + \Omega(\mathbf{W}, \mathbf{R}_{\Phi_0}) + \tilde{\sigma}_z^2 \mathbf{I}_{N_t}$.

As a result, we propose to handle (P) via updating $(\{\mathbf{w}_k\}, \{p_\ell^U\})$ by solving (34) and updating $\{\mathbf{v}_\ell\}$ by (35) in an alternating fashion, as shown in Algorithm 2. If given the initial beamforming vectors $\{\mathbf{v}_\ell^{(0)}\}$, problem (34) (i.e., (33)) is infeasible, then the algorithm shall declare infeasibility. However, this does not necessarily imply that (P) is infeasible. When (33) is feasible given the initial $\{\mathbf{v}_\ell^{(0)}\}$, one can show that the AO algorithm converges to a KKT solution of (P).

Theorem 2 Suppose that (33) is feasible given initial $\mathbf{v}_\ell^{(0)}$, $\ell \in \mathcal{L}$. For Algorithm 2, the total power $\sum_{k=1}^K \|\mathbf{w}_k^{(t)}\|_2^2 + \sum_{\ell=1}^L (p_\ell^U)^{(t)}$ is non-increasing with the iteration number t and converges as $t \rightarrow \infty$. Moreover, any limit point of $(\{\mathbf{w}_k^{(t+1)}\}, \{(p_\ell^U)^{(t+1)}\}, \{\mathbf{v}_\ell^{(t+1)}\})$ is a KKT point of (P).

Proof: It is easy to show that the objective value $\sum_{k=1}^K \|\mathbf{w}_k^{(t)}\|_2^2 + \sum_{\ell=1}^L (p_\ell^U)^{(t)}$ is non-increasing. The proof of $(\{\mathbf{w}_k^{(t+1)}\}, \{(p_\ell^U)^{(t+1)}\}, \{\mathbf{v}_\ell^{(t+1)}\})$ converging to a KKT point of problem (P) can follow [31, Section IV]. Firstly, analogous to [31, Lemma 4] and using (34), one can show that problem (33) has a unique solution up to a phase rotation to $\{\mathbf{w}_k\}$. Then, following a similar argument as in [31, Proposition 1], one can prove that any limit point of $(\{\mathbf{w}_k^{(t+1)}\}, \{(p_\ell^U)^{(t+1)}\}, \{\mathbf{v}_\ell^{(t+1)}\})$ is a KKT point of (P). The details are omitted here. ■

Algorithm 2 Proposed SOCP-based AO algorithm for solving (P) in (20).

- 1: **Given** initial UL beamforming vectors $\mathbf{v}_\ell^{(0)}$, $\ell = 1, \dots, L$; set $t \leftarrow 0$. Declare infeasibility if (34) is infeasible given $\{\mathbf{v}_\ell^{(0)}\}$.
 - 2: **repeat**
 - 3: Given $\{\mathbf{v}_\ell^{(t)}\}$, obtain $(\{\mathbf{w}_k^{(t+1)}\}, \{(p_\ell^U)^{(t+1)}\})$ by solving (34) (i.e., the SOCP reformulation of (33)).
 - 4: Given $(\{\mathbf{w}_k^{(t+1)}\}, \{(p_\ell^U)^{(t+1)}\})$, obtain $\{\mathbf{v}_\ell^{(t+1)}\}$ by (35).
 - 5: $t \leftarrow t + 1$.
 - 6: **until** a predefined convergence condition is satisfied.
-

B. FD UDD and Fixed Point Method for Solving (33)

In this subsection, we show that the FD problem (33) has a duality property that resembles the UDD (i.e., Lemma 3) in HD systems. Moreover, similar to the fixed-point method for solving the HD problems [19], [22], the FD problem (33) can also be solved by efficient fixed-point iterations. To illustrate these methods, let us consider a partial Lagrange dual problem of (33)

$$\max_{\boldsymbol{\nu} \succeq \mathbf{0}} \psi(\boldsymbol{\nu}) + (\sigma_z^2 - \gamma^{\text{ADC}}) \mathbf{1}^T \boldsymbol{\nu} \quad (36)$$

where $\psi(\boldsymbol{\nu})$ is given by

$$(P2) \quad \psi(\boldsymbol{\nu}) = \min_{\{\mathbf{w}_k\}, \{p_\ell^U \geq 0\}} \sum_{k=1}^K \mathbf{w}_k^H \mathbf{B} \mathbf{w}_k + \sum_{\ell=1}^L p_\ell^U b_\ell \quad (37a)$$

$$\text{s.t. (33b), (33c).} \quad (37b)$$

Here, $\boldsymbol{\nu} = [\nu_1, \dots, \nu_{N_t}]^T \in \mathbb{R}^{N_t}$ contain the dual variables associated with the ADC input power constraint (33d); for notational simplicity, it is also defined that $\mathbf{B} \triangleq \mathbf{I}_{N_t} + \sum_{n=1}^{N_t} \nu_n \boldsymbol{\Upsilon}_n(\{\mathbf{R}_{\Phi_0, m}\})$ and $b_\ell \triangleq 1 + \sum_{n=1}^{N_t} \nu_n |\mathbf{e}_n^T \mathbf{g}_j|^2$ for all $\ell \in \mathcal{L}$. Owing to the hidden convexity of problem (33) (i.e., (34)), one can show that problem (33) in fact has a zero duality gap (see, e.g., [32, Proposition 1]). Therefore, solving the dual problem (36) is equivalent to solving problem (33). As a standard approach, one can solve problem (36) by the subgradient method [33]. In each update of the subgradient method, one has to solve the subproblem (P2).

The following proposition shows that (P2) has an equivalent duality problem:

Proposition 4 Suppose that problem (P2) in (37) is feasible. Then (P2) achieves the same optimal objective value as the following problem

$$(\{\tilde{\mathbf{w}}_k^*\}, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \arg \min_{\{\tilde{\mathbf{w}}_k\}, \boldsymbol{\lambda}, \boldsymbol{\mu} \succeq \mathbf{0}} \boldsymbol{\lambda}^T \boldsymbol{\sigma}^2 + \tilde{\sigma}_z^2 \boldsymbol{\mu}^T \mathbf{1} \quad (38a)$$

$$\text{s.t. } \frac{\lambda_k |\tilde{\mathbf{w}}_k^H \mathbf{h}_k|^2 / \rho_k^D}{\tilde{\mathbf{w}}_k^H \mathbf{Q}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \tilde{\mathbf{w}}_k} \geq 1, \quad k \in \mathcal{K}, \quad (38b)$$

$$\frac{\mu_\ell |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L \mu_j \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \sum_{k=1}^K \lambda_k |f_{\ell k}|^2 + b_\ell} \geq 1, \quad \ell \in \mathcal{L}, \quad (38c)$$

$$\|\tilde{\mathbf{w}}_k\|_2 = 1, \quad k \in \mathcal{K}, \quad (38d)$$

where $\boldsymbol{\sigma}^2 \triangleq [\sigma_1^2, \dots, \sigma_K^2]^T$ and $\mathbf{Q}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \sum_{i=1}^K \lambda_i \tilde{\mathbf{H}}_i + \sum_{\ell=1}^L \mu_\ell \boldsymbol{\Lambda}(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) + \mathbf{B}$. Moreover, the optimal beamforming direction of (P2) can be obtained via (38) as

$$\frac{\mathbf{w}_k^*}{\|\mathbf{w}_k^*\|_2} = \tilde{\mathbf{w}}_k^* = \frac{\mathbf{Q}^{-1}(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \mathbf{h}_k}{\|\mathbf{Q}^{-1}(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \mathbf{h}_k\|_2}, \quad k \in \mathcal{K}, \quad (39)$$

where $\{\mathbf{w}_k^*\}$ denotes the optimal $\{\mathbf{w}_k\}$ of (P2).

Proof: See Appendix D. ■

By comparing (38b) and (38c) with (33b) and (33c), respectively, one can see that problem (38) can be regarded as a weighted power minimization problem for a *virtual* FD system with K UMUs and L DMUs. In particular, $\{\mu_\ell\}$ and $\{\lambda_k\}$ are the DL and UL powers in the virtual FD system, respectively, and all the variables and channels originally associated with the DL (resp. UL) are now associated with UL (resp. DL) in the virtual system.

Analogous to the HD UDD which is used to develop a fixed-point method [19], [22], we use the above FD UDD to develop a new fixed-point method to solve (P2). To see this, notice that (38b) and (38c) must hold with equality at the optimum. By substituting (39) into them, constraints (38b) and (38c) can be equivalently written as

$$\lambda_k = \mathcal{F}_k^U(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \frac{1}{\mathbf{h}_k^H \mathbf{Q}^{-1}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \mathbf{h}_k / \rho_k^D}, \quad \forall k \in \mathcal{K}, \quad (40)$$

$$\mu_\ell = \mathcal{F}_\ell^D(\boldsymbol{\lambda}, \boldsymbol{\mu}) \triangleq \frac{\sum_{j=1}^L \mu_j \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \sum_{k=1}^K \lambda_k |f_{\ell k}|^2 + b_\ell}{|\tilde{\mathbf{v}}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}, \quad \forall \ell \in \mathcal{L}. \quad (41)$$

By defining

$$\begin{aligned} \mathcal{F}(\boldsymbol{\lambda}, \boldsymbol{\mu}) &\triangleq [\mathcal{F}_1^U(\boldsymbol{\lambda}, \boldsymbol{\mu}), \dots, \mathcal{F}_K^U(\boldsymbol{\lambda}, \boldsymbol{\mu}), \\ &\quad \mathcal{F}_1^D(\boldsymbol{\lambda}, \boldsymbol{\mu}), \dots, \mathcal{F}_L^D(\boldsymbol{\lambda}, \boldsymbol{\mu})]^T, \end{aligned}$$

we obtain the following fixed point equation

$$[\boldsymbol{\lambda}^T \boldsymbol{\mu}^T]^T = \mathcal{F}(\boldsymbol{\lambda}, \boldsymbol{\mu}). \quad (42)$$

Lemma 4 *Suppose that problem (P2) in (37) is feasible. Then the optimal solution $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ of (38) is the unique fixed point of (42). Moreover, given any initial $(\boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)})$, $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ can be achieved by the following fixed-point iterations*

$$[(\boldsymbol{\lambda}^{(t+1)})^T (\boldsymbol{\mu}^{(t+1)})^T]^T \leftarrow \mathcal{F}(\boldsymbol{\lambda}^{(t)}, \boldsymbol{\mu}^{(t)}) \quad (43)$$

as $t \rightarrow \infty$.

Proof: As shown in (42), once (P2) is feasible, $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is a fixed point. So the fixed point exists for (42). To show that $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is the unique fixed point and the fixed-point iterations in (43) converges to $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ for arbitrary $(\boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)})$, it suffices to show that \mathcal{F} is a standard interference function [34, Theorem 2]; that is, each \mathcal{F}_k^U and \mathcal{F}_ℓ^D should satisfy positivity, monotonicity and scalability properties, for all $k \in \mathcal{K}$ and $\ell \in \mathcal{L}$. The part of \mathcal{F}_k^U can be proved following exactly the same arguments as in [19, Appendix II], while the part of \mathcal{F}_ℓ^D is easy to verify to be true. The details are omitted. ■

Once $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ is obtained by the above fixed-point iterations, the optimal DL beamforming direction $\{\tilde{\mathbf{w}}_k^*\}$ can be obtained by (39). What remains for solving (P2) is to obtain the optimal DL transmission powers $\{p_k^{D*}\}$ and UL transmission power $\{p_\ell^{U*}\}$. To show how they can be obtained, let us first introduce some notations. Let \mathbf{S}_{11} be a $K \times K$ matrix whose k th diagonal entry is $|\mathbf{h}_k^H \tilde{\mathbf{w}}_k^*|^2 / \rho_k^D - (\tilde{\mathbf{w}}_k^*)^H \tilde{\mathbf{H}}_k \tilde{\mathbf{w}}_k^*$ and (k, i) th off-diagonal entry is $-(\tilde{\mathbf{w}}_i^*)^H \tilde{\mathbf{H}}_k \tilde{\mathbf{w}}_i^*$. Moreover, define \mathbf{S}_{12} as a $K \times L$ matrix with the (k, ℓ) th entry being $-|f_{\ell k}|^2$. Similarly, let us define an $L \times L$ matrix \mathbf{S}_{22} which has the ℓ th diagonal entry being $|\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U - \mathbf{v}_\ell^H \tilde{\mathbf{G}}_\ell \mathbf{v}_\ell$ and the (k, i) th off-diagonal entry being $-\mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell$. Lastly, define an $L \times K$ matrix \mathbf{S}_{21} whose (ℓ, k) th entry is $-\tilde{\mathbf{w}}_k^H \mathbf{\Lambda}(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) \tilde{\mathbf{w}}_k$. Then, given $\{\mathbf{w}_k^*\}$, constraints (33b) and (33c) can be compactly expressed as

$$\mathbf{S}(\{\tilde{\mathbf{w}}_k^*\}, \{\mathbf{v}_\ell\}) \begin{bmatrix} \mathbf{p}^D \\ \mathbf{p}^U \end{bmatrix} \succeq \begin{bmatrix} \boldsymbol{\sigma}^2 \\ \tilde{\sigma}_z^2 \mathbf{1} \end{bmatrix}, \quad (44)$$

where $\mathbf{p}^D = [p_1^D, \dots, p_K^D]^T$ and

$$\mathbf{S}(\{\tilde{\mathbf{w}}_k^*\}, \{\mathbf{v}_\ell\}) = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}. \quad (45)$$

The following lemma gives the closed-form solution of optimal power $(\{p_k^{D*}\}, \{p_\ell^{U*}\})$.

Lemma 5 Suppose that problem (P2) in (37) is feasible, and that the optimal DL beamforming directions $\{\tilde{\mathbf{w}}_k^*\}$ are given. Then the optimal DL and UL powers $\mathbf{p}^{\text{D}\star}$ and $\mathbf{p}^{\text{U}\star}$ are uniquely given by

$$\begin{bmatrix} \mathbf{p}^{\text{D}\star} \\ \mathbf{p}^{\text{U}\star} \end{bmatrix} = \mathbf{S}^{-1}(\{\tilde{\mathbf{w}}_k^*\}, \{\mathbf{v}_\ell\}) \begin{bmatrix} \sigma^2 \\ \tilde{\sigma}_z^2 \mathbf{1} \end{bmatrix} \succ \mathbf{0}. \quad (46)$$

Proof: Given that (P2) is feasible, constraints (33b) and (33c) will hold with equality at the optimum and $\mathbf{p}^{\text{D}\star}$ and $\mathbf{p}^{\text{U}\star}$ will be a solution to the linear system (44). To show that $\mathbf{p}^{\text{D}\star}$ and $\mathbf{p}^{\text{U}\star}$ are the unique solution, one can follow a similar arguments as in [25, Lemma 1]. A sufficient and necessary condition for the proof to be valid is that the diagonal elements of \mathbf{S}_{11} and \mathbf{S}_{22} have to be positive, i.e., $|\mathbf{h}_k^H \tilde{\mathbf{w}}_k^*|^2 / \rho_k^{\text{D}} - (\tilde{\mathbf{w}}_k^*)^H \tilde{\mathbf{H}}_k \tilde{\mathbf{w}}_k^* > 0$ for all $k \in \mathcal{K}$ and $|\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^{\text{U}} - \mathbf{v}_\ell^H \tilde{\mathbf{G}}_\ell \mathbf{v}_\ell > 0$ for all $\ell \in \mathcal{L}$. Recall from Proposition 4 that the optimal DL beamforming direction $\{\mathbf{w}_k^*\}$ must satisfy (38b), i.e.,

$$\begin{aligned} & \lambda_k^* |\mathbf{h}_k^H \tilde{\mathbf{w}}_k^*|^2 / \rho_k^{\text{D}} - (\tilde{\mathbf{w}}_k^*)^H (\lambda_k^* \tilde{\mathbf{H}}_k) \tilde{\mathbf{w}}_k^* \\ & - (\mathbf{w}_k^*)^H \left(\sum_{i \neq k}^K \lambda_i^* \tilde{\mathbf{H}}_i + \sum_{\ell=1}^L \mu_\ell^* \Lambda(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) + \mathbf{B} \right) \tilde{\mathbf{w}}_k^* \geq 0. \end{aligned}$$

Since $\sum_{i \neq k}^K \lambda_i^* \tilde{\mathbf{H}}_i + \sum_{\ell=1}^L \mu_\ell^* \Lambda(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) + \mathbf{B} \succ \mathbf{0}$ (due to $\mathbf{B} \succ \mathbf{0}$), we have $|\mathbf{h}_k^H \tilde{\mathbf{w}}_k^*|^2 / \rho_k^{\text{D}} - (\tilde{\mathbf{w}}_k^*)^H \tilde{\mathbf{H}}_k \tilde{\mathbf{w}}_k^* > 0$. Analogously, since $\{\mathbf{v}_\ell\}$ satisfies (38c) and $\sum_{j \neq \ell}^L \mu_j^* \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \sum_{k=1}^K \lambda_k^* |f_{\ell k}|^2 + b_\ell > 0$, it is clear that $|\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^{\text{U}} - \mathbf{v}_\ell^H \tilde{\mathbf{G}}_\ell \mathbf{v}_\ell > 0$. ■

It is worth noting that Proposition 4, Lemma 4 and Lemma 5 essentially generalize the existing UDD and fixed-point method in HD systems [22], [25], [27] to the FD system. As seen, this generalization enables an efficient way to solve (P2). By combining it with the subgradient method, we come up with a low-complexity algorithm to solve problem (33) (i.e., (36)), which is shown in Algorithm 3. In particular, step 3 to step 7 of Algorithm 3 are the fixed-point iterations to obtain $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$; step 8 and step 9 are based on (39) and Lemma 5; step 10 is the subgradient update for dealing with the ADC input power constraint (33d), where $s^{(r)}$ is the step size¹. Comparing to directly solving the SOCP (34) using a general purpose solver, the proposed fixed-point based algorithm in Algorithm 3 is computationally more efficient, and therefore improving the computational efficiency of Algorithm 2.

V. SIMULATION RESULTS

In the simulation, we consider a wireless system as described in Section II-A and Figure 1. The FD BS has 10 antennas ($N_t = 10$) for simultaneous UL and DL communications [8], [24]. The channel

¹Note that if the ADC input power constraint (33d) is not considered, then the sbgradient loop (i.e., Steps 1, 2 and 10 to 12) can be removed; in that case, \mathbf{B} and b_ℓ 's reduce to \mathbf{I}_{N_t} and 1, respectively.

Algorithm 3 Proposed fixed-point based algorithm for solving subproblem (33) in Algorithm 2.

- 1: **Given** initial variables $\boldsymbol{\nu}^{(0)} = \mathbf{0}$; set $r \leftarrow 0$.
 - 2: **repeat**
 - 3: set $t \leftarrow 0$ and initial $\boldsymbol{\lambda}^{(0)} = \mathbf{0}$ and $\boldsymbol{\mu}^{(0)} = \mathbf{0}$.
 - 4: **repeat**
 - 5: $[\mathbf{u}^{(t+1)} := (\boldsymbol{\lambda}^{(t+1)})^T (\boldsymbol{\mu}^{(t+1)})^T]^T \leftarrow \mathcal{F}(\boldsymbol{\lambda}^{(t)}, \boldsymbol{\mu}^{(t)})$
 - 6: $t \leftarrow t + 1$
 - 7: **until** $\|\mathbf{u}^{(t+1)} - \mathbf{u}^{(t)}\|_2 / \|\mathbf{u}^{(t)}\|_2 \leq \epsilon_1$. Denote the converged results as $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$.
 - 8: Obtain $\{\tilde{\mathbf{w}}_k^*\}$ by (39), and obtain $\{(p_k^{\text{D}^*})\}$ and $\{(p_\ell^{\text{U}^*})\}$ by (46).
 - 9: $\mathbf{w}_k^* \leftarrow \sqrt{(p_k^{\text{D}^*})} \tilde{\mathbf{w}}_k^*, \forall k \in \mathcal{K}$.
 - 10: $\boldsymbol{\nu}^{(r+1)} \leftarrow \boldsymbol{\nu}^{(r)} + s^{(r)} (\text{diag}(\text{ADC}(\mathbf{W}^*, \mathbf{p}^{\text{U}^*}) - \gamma^{\text{ADC}} \mathbf{I}_{N_t}))$.
 - 11: $r \leftarrow r + 1$.
 - 12: **until** $\|\boldsymbol{\nu}^{(r+1)} - \boldsymbol{\nu}^{(r)}\|_2 / \|\boldsymbol{\nu}^{(r)}\|_2 \leq \epsilon_2$
-

coefficients $\{\mathbf{h}_i\}, \{\mathbf{g}_j\}, \{f_{ji}\}$ and \mathbf{H}_0 are composed by large-scale path loss as well as small scale Rayleigh fadings. In particular, the path loss between the MUs and the BS is set to -80 dB and that between the UMUs and DMUs is set to -83 dB. The SI channel has a -10 dB path loss [8], [24]. Besides, following [8], there is additional -24 dB cross-talk path loss for neighboring antennas and further -6 dB path loss for farther antennas. That is, the path loss between (transmit) antenna i and (receive) antenna j in the SI channel is -10 dB for all $i = j$ and $-34 - 6|i - j - 1|$ dB, for all $i \neq j$, where $i, j \in \{1, \dots, N_t\}$. Define \mathbf{T} (resp. \mathbf{C}) as an N_t by N_t Toeplitz matrix with the first row being $[1, 10^{-\frac{24}{20}}, 10^{-\frac{30}{20}}, \dots, 10^{-\frac{24-6(N_t-2)}{20}}]$ (resp. $[1, 0.9, 0.9^2, \dots, 0.9^{N_t-1}]$). We model the correlation matrix of SI channel \mathbf{H}_0 as

$$\begin{aligned} \mathbf{R}_{\mathbf{H}_0} &= \mathbb{E}[\text{vec}(\mathbf{H}_0) \text{vec}(\mathbf{H}_0)^T] \\ &= \left(10^{-\frac{10}{10}} [\text{vec}(\mathbf{T}) \text{vec}(\mathbf{T})^T] \right) \odot \left(\mathbf{1}_{N_t \times N_t} \otimes \mathbf{C} \right), \end{aligned} \quad (47)$$

where \odot is the (element-wise) Hadamard product; \otimes is the Kronecker product and $\mathbf{1}_{N_t \times N_t}$ is the N_t by N_t all-one matrix. The first term in the RHS of (47) accounts for the cross-talk path loss, while the second term is for modeling the decreasing correlation between adjacent antennas. To model \mathbf{R}_{Φ_0} , we assume that the analog SIC scheme uses pilot-aided linear minimum mean squared error (LMMSE) channel estimation to estimate \mathbf{H}_0 [35]. Then, \mathbf{R}_{Φ_0} is given by $\mathbf{R}_{\Phi_0} = \mathbf{R}_{\mathbf{H}_0} - \mathbf{R}_{\mathbf{H}_0} (\mathbf{R}_{\mathbf{H}_0} + \frac{\sigma_z^2}{E} \mathbf{I}_{N_t^2})^{-1} \mathbf{R}_{\mathbf{H}_0}$,

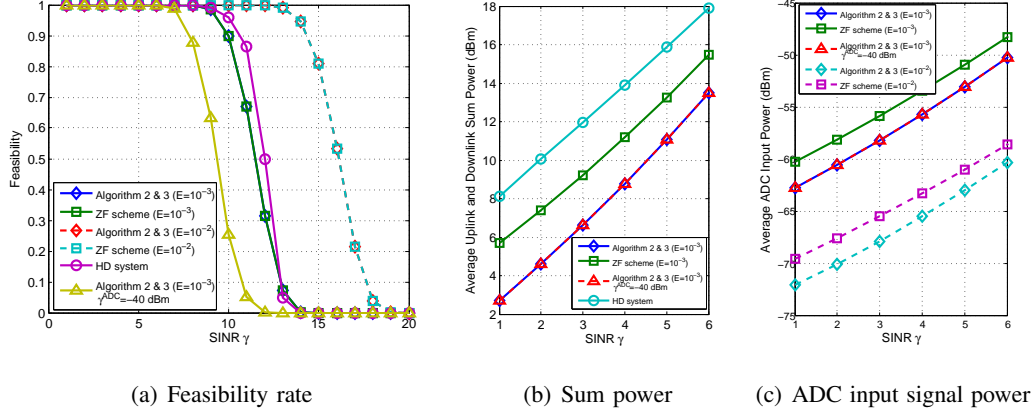


Fig. 4: Performance comparison of the proposed Algorithm 2 & 3 and the ZF scheme; $N_t = 10$, $K = L = 8$.

where $E > 0$ denotes the energy of training signals. As seen, the larger E is, the more powerful the analog SIC is. If not mentioned specifically, we set various parameters as follows: $\sigma_z^2 = \sigma_i^2 = -85$ dBm, $\beta_1 = \beta_2 = -30$ dB, $\delta_1 = -50$ dB and $\delta_2 = -20$ dB. For Algorithm 1, P_{\max} is set to 40 dBm and ϵ is set to 10^{-3} . Problems (29) and (30) are solved by the classical fixed-point method [19], [22]. For Algorithm 2, the stopping condition is set to the relative improvement of objective value of (P) being less than 10^{-6} . Instead of using a convex solver to solve the SOCP (34), we use Algorithm 3 to solve problem (33) since they two yield the same performance theoretically. The parameters are set to $\epsilon_1 = 10^{-9}$, $\epsilon_2 = 10^{-3}$ and $s^{(r)} = 1$ (constant step size). The initial $\{v_\ell^{(0)}\}$ is set to the zero forcing (ZF) beamformer, i.e., $v_\ell^{(0)} = G^\dagger e_\ell$, for all $\ell \in \mathcal{L}$, where G^\dagger is the pseudo inverse of $G = [g_1, \dots, g_L]^T$. Note that if the AO algorithm runs only one iteration, then it is the same as the ZF based scheme in [17]. Besides, if not mentioned specifically, the ADC input power constraint (20d) is not considered in order to assess how the ADC input signal power varies with system parameters if unconstrained.

For simplicity, we let all UMUs and DMUs have the same SINR requirement, i.e., $\gamma \triangleq \gamma_i^D = \gamma_\ell^U$ for all $i \in \mathcal{K}$ and $\ell \in \mathcal{L}$. For a fair comparison, we set the SINR of MUs in the HD system as $\gamma_{\text{HD}} \triangleq 2^{2 \log_2(1+\gamma)} - 1$. This implies that the information rate achieved by the HD MUs should be twice of that by the FD MUs [17]. All simulation results are obtained by averaging over 500 channel realizations.

Example 1: In Figure 4, we display the simulation results by comparing the proposed AO algorithm (i.e., Algorithm 2 and 3) with the ZF scheme and the HD system. The number of DMUs and UMUs are eight ($K = L = 8$). First of all, one can observe from Figure 4(a) that the proposed AO algorithm

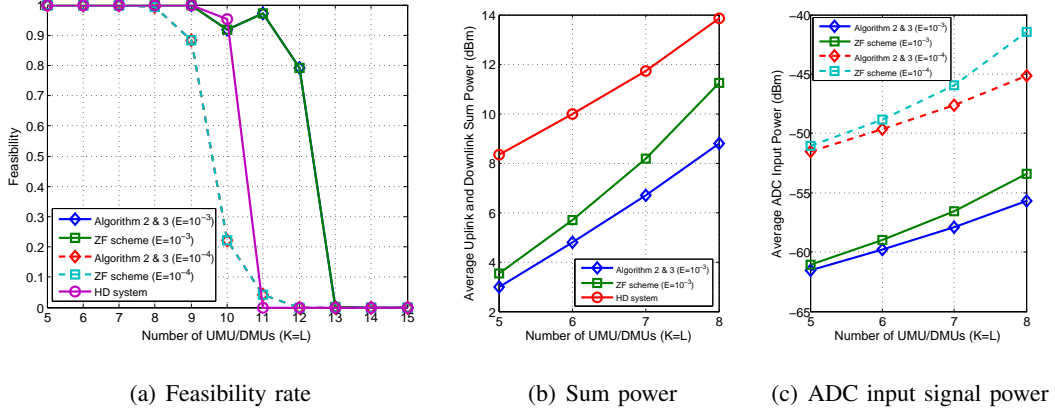


Fig. 5: Performance comparison of the proposed Algorithm 2 & 3 and the ZF scheme; $N_t = 10$, $\gamma = 5$ dB.

has the same feasibility rate as the ZF scheme, which is expected, since the proposed AO algorithm is initialized by the ZF receive beamformer. However, one can see from Figure 4(b) and Figure 4(c) that² the AO algorithm can yield about 3 dB lower sum power and ADC input power than the ZF scheme. It can also be seen from the two figures that for both $E = 10^{-3}$ and $E = 10^{-2}$, the FD system using the proposed AO algorithm is more power efficient than the HD system when both systems are feasible.

In Figure 4, we also present the performance of the proposed AO algorithm when the ADC input power constraint (20d) is imposed with $\gamma^{\text{ADC}} = -40$ dBm. One can see from Figure 4(a) that the feasibility rate drops compared to that without the ADC input power constraint for $\gamma \geq 7$ dB. This implies that there exist realizations for which the ADC input signal power is higher than -40 dBm and this happens more frequently when γ increases. Since from Figure 4(c) that the ADC input power is less than -40 dBm for $\gamma \leq 6$ dB, Figure 4(b) shows that the AO algorithm with the ADC input power constraint performs equally well as its counterpart without the ADC input power constraint in this regime.

Example 2: In Figure 5, we present the results by considering various numbers of DMU/UMUs in the network. The SINR requirement is set to $\gamma = 5$ dB. As expected, the system performance (feasibility rate and sum power) deteriorates when the number of MUs in the network increases. The ADC input power also increases since the SI power can increase both with the number of UMUs and the number of DMUs. One may notice that the feasibility rate of the proposed AO algorithm ($E = 10^{-3}$) oscillates between $K = L = 10$ and $K = L = 11$. This is because the AO algorithm is initialized by different

²For fair comparison, in Figure 4(b) and Figure 4(c), we only show the results for which all schemes under test are 100% feasible, i.e., from $\gamma = 1$ to $\gamma = 6$. This principle also applies to Figure 5 and Figure 6.

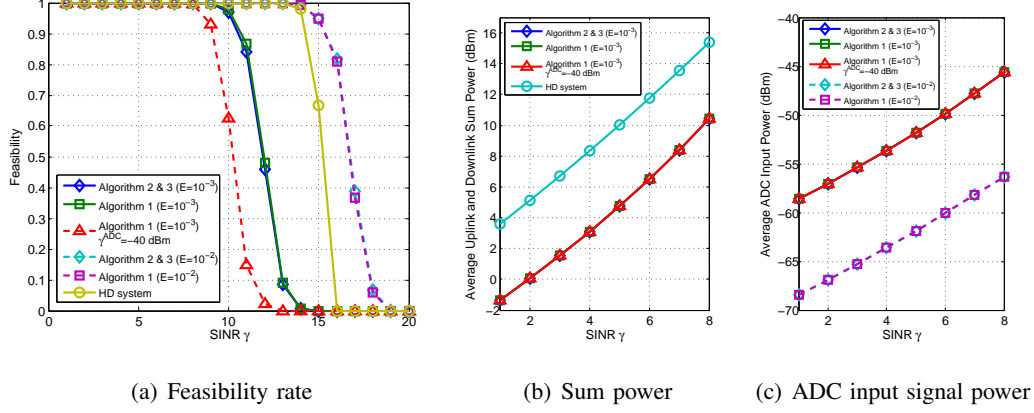


Fig. 6: Performance of the proposed Algorithm 1, 2 and 3; $N_t = 10$, $K = L = 5$, $\mathbf{R}_{\Phi_0} = \sigma_{\Phi_0}^2 \mathbf{I}_{N_t^2}$ where $\sigma_{\Phi_0}^2 = -10$ dB.

$\{\mathbf{v}_\ell^{(0)}\}$ when the number of UMUs changes. Finally, it can be observed from Figure 5(b) and Figure 5(c) that the more number of MUs in the network, the better the AO algorithm performs than the ZF scheme.

Example 3: In the last example, let us examine the performance of Algorithm 1 by assuming that the SI channel errors are i.i.d. Specifically, we let $\mathbf{R}_{H_0} = \sigma_{H_0}^2 \mathbf{I}_{N_t^2}$ where $\sigma_{H_0}^2 = -10$ dB³, and let $K = L = 5$. The simulation results are presented in Figure 6. Firstly, one can observe from Figure 6(a) that there exists small discrepancy between the feasibility rates of the AO algorithm (Algorithm 2 and 3) and the bisection algorithm (Algorithm 1), especially when the feasibility rates are not 100%. We suspect that the AO algorithm actually achieves the same solution as the bisection algorithm under the simulation setting, and the discrepancy in the feasibility rate are caused by some numerical issues. This is evidenced by observing from Figure 6(b) and Figure 6(c) that the AO and bisection algorithms essentially yield the same sum power and ADC input power when both methods are feasible. Therefore, the simulation results imply that the AO algorithm may have achieved optimal or near-optimal solutions for the cases with $\mathbf{R}_{H_0} = \sigma_{H_0}^2 \mathbf{I}_{N_t^2}$. From Figure 6(a), we again see that with the ADC input power constraint, the feasibility rates of the considered design problem decreases.

VI. CONCLUSION

In this paper, by taking into account the non-ideal RF chains and analog/digital SIC, we have formulated the QoS-based linear transceiver design problem (P) which is not only constrained by the minimum SINR

³So $\mathbf{R}_{\Phi_0} = \sigma_{\Phi_0}^2 \mathbf{I}_{N_t^2}$ where $\sigma_{\Phi_0}^2$ is approximately -95 dB when $E = 10^{-3}$ and -105 dB when $E = 10^{-2}$.

requirement of MUs but also by the maximum power at the ADC input. While such problem is non-convex and difficult to solve in general, we have shown that it can be solved to the global optimum when the SI channel estimation errors are i.i.d. Specifically, we have developed a bisection algorithm (Algorithm 1) that can achieve the global optimal solution in a polynomial-time complexity. To handle (P) in general, we have proposed a suboptimal AO algorithm (Algorithm 2). Moreover, we have generalized the UDD in the HD system to the FD system and proposed an efficient fixed-point based algorithm for solving (33) (Algorithm 3). The presented simulation results have shown that the proposed AO algorithm outperforms the ZF scheme. Moreover, when the SI channel estimation errors are i.i.d., the simulation results also suggest that the AO algorithm achieves a near-optimal solution. It is also observed that the FD system is more power efficient than the HD system especially when the QoS requirements are less stringent or when the number of MUs is moderate.

APPENDIX A

DERIVATION OF (12) AND (13)

The derivation involves basic vector and matrix algebra. Note that we can write the first term of in the RHS of (11) as

$$\begin{aligned} \delta_1 \mathbf{v}_\ell^H \mathbb{E}[\Phi_0 \mathbf{x}(t) \mathbf{x}^H(t) \Phi_0^H] \mathbf{v}_\ell \\ = \delta_1 \mathbf{v}_\ell^H \mathbb{E} \left[\Phi_0 \mathbf{W} \mathbf{W}^H \Phi_0^H \right] \mathbf{v}_\ell \end{aligned} \quad (\text{A.1})$$

$$= \delta_1 \sum_{k=1}^K \mathbf{w}_k^H \mathbb{E} \left[\Phi_0^H \mathbf{v}_\ell \mathbf{v}_\ell^H \Phi_0 \right] \mathbf{w}_k. \quad (\text{A.2})$$

Besides, by (1), we can write the second term in the RHS of (11) as

$$\begin{aligned} \delta_2 \mathbf{v}_\ell^H \mathbb{E}[\Phi_0 \mathbf{u}_{\text{tx}}[n] \mathbf{u}_{\text{tx}}^H[n] \Phi_0^H] \mathbf{v}_\ell \\ = (\delta_2 \beta_1) \mathbf{v}_\ell^H \mathbb{E}[\Phi_0 \text{diag}(\mathbf{W} \mathbf{W}^H) \Phi_0^H] \mathbf{v}_\ell \end{aligned} \quad (\text{A.3})$$

$$\begin{aligned} = (\delta_2 \beta_1) \sum_{k=1}^K \sum_{n=1}^{N_t} (\mathbf{v}_\ell^H \mathbb{E}[\Phi_0 \mathbf{e}_n \mathbf{e}_n^T \Phi_0^H] \mathbf{v}_\ell) |w_{kn}|^2 \\ = (\delta_2 \beta_1) \sum_{k=1}^K \mathbf{w}_k^H \text{diag} \left(\{ \mathbf{v}_\ell^H \mathbb{E}[\Phi_0 \mathbf{e}_n \mathbf{e}_n^T \Phi_0^H] \mathbf{v}_\ell \}_n \right) \mathbf{w}_k. \end{aligned} \quad (\text{A.4})$$

Analogously, by (6) and (8), the third term in the RHS of (11) can be written as

$$\begin{aligned}
& \delta_2 \mathbf{v}_\ell^H \mathbb{E}[\mathbf{u}_{\text{rx}}[n] \mathbf{u}_{\text{rx}}^H[n]] \mathbf{v}_\ell \\
&= (\delta_2 \beta_2) \mathbf{v}_\ell^H \text{diag} \left(\sum_{j=1}^L \mathbf{g}_j \mathbf{g}_j^H p_j^{\text{U}} + \mathbb{E}[\Phi_0 \mathbf{W} \mathbf{W}^H \Phi_0^H] + \beta_1 \mathbb{E}[\Phi_0 \text{diag}(\mathbf{W} \mathbf{W}^H) \Phi_0^H] + \sigma_z^2 \mathbf{I}_{N_t} \right) \mathbf{v}_\ell \\
&= (\delta_2 \beta_2) \mathbf{v}_\ell^H \left(\sum_{j=1}^L p_j^{\text{U}} \text{diag}(|\mathbf{g}_j|^2) \right) \mathbf{v}_\ell + (\delta_2 \beta_2) \mathbf{v}_\ell^H \text{diag} \left(\mathbb{E}[\Phi_0 \mathbf{W} \mathbf{W}^H \Phi_0^H] \right. \\
&\quad \left. + \beta_1 \mathbb{E}[\Phi_0 \text{diag}(\mathbf{W} \mathbf{W}^H) \Phi_0^H] + \sigma_z^2 \mathbf{I}_{N_t} \right) \mathbf{v}_\ell \tag{A.5}
\end{aligned}$$

$$\begin{aligned}
&= (\delta_2 \beta_2) \left[\mathbf{v}_\ell^H \left(\sum_{j=1}^L p_j^{\text{U}} \text{diag}(|\mathbf{g}_j|^2) \right) \mathbf{v}_\ell + \sum_{k=1}^K \mathbf{w}_k^H \left(\sum_{n=1}^{N_t} |v_{\ell,n}|^2 \mathbb{E}\{\Phi_0^H \mathbf{e}_n \mathbf{e}_n^T \Phi_0\} \right) \mathbf{w}_k \right. \\
&\quad \left. + \beta_1 \sum_{k=1}^K \mathbf{w}_k^H \text{diag} \left(\left\{ \mathbf{v}_\ell^H \text{diag} \left(\mathbb{E}[\Phi_0 \mathbf{e}_n \mathbf{e}_n^T \Phi_0^H] \right) \mathbf{v}_\ell \right\}_n \right) \mathbf{w}_k + \sigma_z^2 \|\mathbf{v}_\ell\|_2^2 \right]. \tag{A.6}
\end{aligned}$$

To simplify the notations, define

$$\begin{aligned}
\Omega(\mathbf{W}, \mathbf{R}_{\Phi_0}) &\triangleq \delta_1 \mathbb{E} \left[\Phi_0 \mathbf{W} \mathbf{W}^H \Phi_0^H \right] + (\delta_2 \beta_1) \mathbb{E}[\Phi_0 \text{diag}(\mathbf{W} \mathbf{W}^H) \Phi_0^H] \\
&\quad + (\delta_2 \beta_2) \text{diag} \left(\mathbb{E}[\Phi_0 \mathbf{W} \mathbf{W}^H \Phi_0^H] + \beta_1 \mathbb{E}[\Phi_0 \text{diag}(\mathbf{W} \mathbf{W}^H) \Phi_0^H] + \sigma_z^2 \mathbf{I}_{N_t} \right), \tag{A.7}
\end{aligned}$$

$$\Lambda(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) \triangleq \delta_1 \mathbb{E}\{\Phi_0^H \mathbf{v}_\ell \mathbf{v}_\ell^H \Phi_0\} + \delta_2 \beta_2 \sum_{n=1}^{N_t} |v_{\ell,n}|^2 \bar{\mathbf{R}}_{\Phi_0,n} + \delta_2 \beta_1 \text{diag}(\{\mathbf{v}_\ell^H \tilde{\mathbf{R}}_{\Phi_0,n} \mathbf{v}_\ell\}_n), \tag{A.8}$$

where

$$\mathbf{R}_{\Phi_0,n} \triangleq \mathbb{E}\{\Phi_0 \mathbf{e}_n \mathbf{e}_n^T \Phi_0^H\}, \bar{\mathbf{R}}_{\Phi_0,n} \triangleq \mathbb{E}\{\Phi_0^H \mathbf{e}_n \mathbf{e}_n^T \Phi_0\}, \tag{A.9}$$

$$\tilde{\mathbf{R}}_{\Phi_0,n} \triangleq \mathbf{R}_{\Phi_0,n} + \beta_2 \text{diag}(\mathbf{R}_{\Phi_0,n}), \tag{A.10}$$

and $\tilde{\mathbf{G}}_j \triangleq \mathbf{g}_j \mathbf{g}_j^H + \delta_2 \beta_2 \text{diag}(|\mathbf{g}_j|^2)$ and $\tilde{\sigma}_z^2 \triangleq (1 + \delta_2 \beta_2) \sigma_z^2$. Then, (12) and (13) can be obtained from (A.1) to (A.8).

APPENDIX B

PROOF OF PROPOSITION 2

To show part (a), suppose that (P1) has two sets of solutions with corresponding $\{\mathbf{w}_k, p_\ell^{\text{U}}\}$ denoted by $\{\hat{\mathbf{w}}_k, \hat{p}_\ell^{\text{U}}\}$ and $\{\tilde{\mathbf{w}}_k, \tilde{p}_\ell^{\text{U}}\}$. Let $\hat{\eta} = \sum_{k=1}^K \|\hat{\mathbf{w}}_k\|_2^2$ and $\tilde{\eta} = \sum_{k=1}^K \|\tilde{\mathbf{w}}_k\|_2^2$. Then it should be

$$\hat{\eta} + \sum_{\ell=1}^L \hat{p}_\ell^{\text{U}} = \tilde{\eta} + \sum_{\ell=1}^L \tilde{p}_\ell^{\text{U}} \tag{A.11}$$

Without loss of generality, suppose that $\hat{\eta} < \tilde{\eta}$. Then $\sum_{\ell=1}^L \hat{p}_\ell^U > \sum_{\ell=1}^L \tilde{p}_\ell^U$. According to the proof of Proposition 1, given a value of $\sum_{k=1}^K \|\mathbf{w}_k\|_2^2 = \eta$, the optimal $\{p_\ell^U\}$ must satisfy (32) and is unique. As a result, for $\hat{\eta} < \tilde{\eta}$, we must have $\sum_{\ell=1}^L \hat{p}_\ell^U < \sum_{\ell=1}^L \tilde{p}_\ell^U$ which is a contradiction. Therefore, we obtain $\hat{\eta} = \tilde{\eta}$.

To show part (b), we observe that when $\eta = \eta^* = \sum_{k=1}^K \|\mathbf{w}_k^*\|_2^2$, the UL powers $\{p_\ell^U\}$ of both (P1) and (P $_\eta$) are uniquely determined by the system equations in (32). Therefore, by Lemma 2, given $\eta = \eta^*$, (P $_\eta$) has the same set of optimal UL powers and UL beamformers as (P1), which we denote as $\{p_\ell^{U*}\}$ and $\{\mathbf{v}_\ell^*\}$, respectively. With $\{p_\ell^U, \mathbf{v}_\ell\}$ fixed by $\{p_\ell^{U*}, \mathbf{v}_\ell^*\}$ in (P1) and (P $_\eta$), one can show that the optimal DL beamformers $\{\mathbf{w}_k\}$ for both problems (P1) and (P $_\eta$) must also be solutions to the following problem

$$\min_{\{\mathbf{w}_k\}} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \quad (\text{A.12a})$$

$$\text{s.t. } \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2 / \rho_i^D}{\sum_{k=1}^K \mathbf{w}_k^H \tilde{\mathbf{H}}_i \mathbf{w}_k + \hat{\sigma}_i^2(\mathbf{p}^{U*})} \geq 1, \quad i \in \mathcal{K}, \quad (\text{A.12b})$$

$$\sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \leq \eta^*, \quad (\text{A.12c})$$

$$\overline{\text{ADC}}(\mathbf{W}, \mathbf{p}^{U*}) \preceq \gamma^{\text{ADC}} \mathbf{I}_{N_t}. \quad (\text{A.12d})$$

It is not difficult to verify that problem (A.12) has a unique solution up to a phase shift (e.g., see [31, Lemma 4]). So $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 = \sum_{k=1}^K \|\mathbf{w}_k^*\|_2^2 = \eta^*$ and therefore $\{\mathbf{w}_k(\eta^*)\}$ of (P $_\eta$) is also optimal to (P1).

Part (c) is true as $\{p_\ell^U(\eta)\}$ increases with η and consequently $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ is increasing with η . ■

APPENDIX C

PROOF OF PROPOSITION 3

Proof of part (a): It is easy to see that $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ is increasing with η . To show the concavity, we first prove that the CCI $\sum_{j=1}^L p_j^U(\eta) |f_{ji}|^2$ in (30b) is a concave function of η for any $i \in \mathcal{K}$. Recall Proposition 1 that the optimal $\{p_\ell^U(\eta)\}$ can be obtained by solving (29). Alternatively, let us consider

the following problem

$$\min_{\{\mathbf{v}_\ell\}, \{p_\ell^U \geq 0\}} \sum_{\ell=1}^L a_\ell p_\ell^U \quad (\text{A.13a})$$

$$\text{s.t.} \quad \frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + (\xi\eta + \tilde{\sigma}_z^2)} \geq 1, \quad \ell \in \mathcal{L}, \quad (\text{A.13b})$$

$$\|\mathbf{v}_\ell\|_2 = 1, \quad \ell \in \mathcal{L}. \quad (\text{A.13c})$$

where $a_\ell > 0$, $\ell = 1, \dots, L$, are some weighting coefficients. Since at the optimum the constraint (A.13b) holds with equality, the optimal $\{p_\ell^U\}$ of problem (A.13) also satisfies (32). As (32) admits only a unique solution, $\{p_\ell^U(\eta)\}$ of (29) is also the optimal solution to problem (A.13). By this fact and by applying Lemma 1 to (A.13), we obtain

$$\sum_{\ell=1}^L a_\ell p_\ell^U(\eta) = \max_{\{p_\ell^U \geq 0\}} \sum_{\ell=1}^L a_\ell p_\ell^U \quad (\text{A.14a})$$

$$\text{s.t.} \quad \sum_{j=1}^L p_j^U \tilde{\mathbf{G}}_j + (\xi\eta + \tilde{\sigma}_z^2) \mathbf{I}_{N_t} \succeq \frac{p_\ell^U}{\rho_\ell^U} \mathbf{g}_\ell \mathbf{g}_\ell^H, \quad \forall \ell \in \mathcal{L}. \quad (\text{A.14b})$$

Define an indicator function as

$$\mathcal{I}(\mathbf{p}^U, \eta) = \begin{cases} 0 & \text{if } (\mathbf{p}^U, \eta) \text{ satisfies (A.14b),} \\ \infty & \text{otherwise.} \end{cases} \quad (\text{A.15})$$

Thus one can write (A.14) as

$$\sum_{\ell=1}^L a_\ell p_\ell^U(\eta) = \max_{\{p_\ell^U \geq 0\}} \sum_{\ell=1}^L a_\ell p_\ell^U - \mathcal{I}(\mathbf{p}^U, \eta). \quad (\text{A.16})$$

Note that $\mathcal{I}(\mathbf{p}^U, \eta)$ is jointly convex w.r.t. η and \mathbf{p}^U . So, by applying the maximization property of concave functions [36, Chapter 3] to (A.16), we obtain that $\sum_{\ell=1}^L a_\ell p_\ell^U(\eta)$ is a concave function of η . By letting $a_\ell = |f_{\ell i}|^2$, we obtain that $\sum_{j=1}^L p_j^U(\eta) |f_{ji}|^2$ in (30b) is a concave function of η .

We use the concavity of $\sum_{j=1}^L p_j^U(\eta) |f_{ji}|^2$ to show that $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ of (\mathbf{P}_η) is also concave. Firstly, by Proposition 1, $\{\mathbf{w}_k(\eta)\}$ can be obtained by solving (30). Since ADC power constraint (30c) can be explicitly written as

$$\sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \leq \frac{\gamma^{\text{ADC}} - \sum_{j=1}^L p_j^U(\eta) |\mathbf{e}_n^T \mathbf{g}_j|^2 - \sigma_z^2}{\sigma_{\Phi_0}^2 (1 + \beta_1)} \quad (\text{A.17})$$

for $n = 1, \dots, N_t$, problem (30) can be solved by first solving

$$\min_{\{\mathbf{w}_k\}} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \quad (\text{A.18a})$$

$$\text{s.t.} \quad \frac{|\mathbf{h}_i^H \mathbf{w}_i|^2 / \rho_i^D}{\sum_{k=1}^K \mathbf{w}_k^H \tilde{\mathbf{H}}_i \mathbf{w}_k + \hat{\sigma}_i^2(\mathbf{p}^U(\eta))} \geq 1, \quad i \in \mathcal{K}, \quad (\text{A.18b})$$

followed by checking whether the optimal $\{\mathbf{w}_k\}$ satisfies (A.17) or not. So given that (\mathbf{P}_η) is feasible, the optimal value of (A.18) is $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$. By applying Lemma 3, (A.18) has a virtual UL problem

$$\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 = \min_{\{\tilde{\mathbf{w}}_i\}, \{\lambda_i \geq 0\}} \sum_{i=1}^K \lambda_i \hat{\sigma}_i^2(\mathbf{p}^U(\eta)) \quad (\text{A.19a})$$

$$\text{s.t. } \frac{\lambda_i |\tilde{\mathbf{w}}_i^H \mathbf{h}_i|^2 / \rho_i^D}{\sum_{k=1}^K \lambda_k \tilde{\mathbf{w}}_i^H \tilde{\mathbf{H}}_k \tilde{\mathbf{w}}_i + \|\tilde{\mathbf{w}}_i\|_2^2} \geq 1, \quad (\text{A.19b})$$

$$\|\tilde{\mathbf{w}}_i\|_2 = 1, \quad i \in \mathcal{K}, \quad (\text{A.19c})$$

which has the same optimal value as (A.18). Notice that, similar to Lemma 2, the optimal $\{\lambda_i\}$ of (A.19) is uniquely determined by equations

$$\frac{\lambda_i}{\rho_i^D} \mathbf{h}_i^H \left(\sum_{k=1}^K \lambda_k \tilde{\mathbf{H}}_k + \mathbf{I}_{N_t} \right)^{-1} \mathbf{h}_i = 1, \quad \forall i \in \mathcal{K}. \quad (\text{A.20})$$

Note that (A.20) is independent of η , and therefore the optimal $\{\lambda_i\}$ of (A.19) is a constant w.r.t. η . Since $\sum_{j=1}^L p_j^U(\eta) |f_{ji}|^2$ is a concave function of η , we conclude that the optimal objective value $\sum_{i=1}^K \lambda_i (\sum_{j=1}^L p_j^U(\eta) |f_{ji}|^2 + \sigma_i^2)$ of (A.19), which is equal to $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$, is a concave function of η . The proof is thus complete.

Proof of part (b): To show sufficiency of part (b), by Proposition 2(c), $F(\eta) = \sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 + \sum_{\ell}^L p_\ell^U(\eta) < F(\eta^*)$ for $\eta < \eta^*$. Suppose that $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 \leq \eta$. Then $\{\mathbf{w}_k(\eta), \mathbf{v}_\ell(\eta), p_\ell^U(\eta)\}$ of (\mathbf{P}_η) is also a feasible solution to $(\mathbf{P1})$. It implies $F(\eta^*) \leq F(\eta)$ which however is a contradiction. So the sufficiency of part (b) is true.

We draw the function $y = \sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ and $y = \eta$ in Figure 3. Specifically, note that $\sum_{k=1}^K \|\mathbf{w}_k(0)\|_2^2 > 0$, and by Proposition 2(b), the function $y = \sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ intersects with the line $y = \eta$ at η^* . Moreover, by part (a), $y = \sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ is concave and increasing. Therefore, $y = \sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ must be below $y = \eta$ when $\eta > \eta^*$; that is, $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 < \eta$ for $\eta > \eta^*$. So the necessity of part (b) is true.

Finally, let us show that $(\mathbf{P1})$ is infeasible if and only if (\mathbf{P}_η) has $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 > \eta$ for all $\eta \in [0, \eta_{\max}]$. The sufficiency is true since by Proposition 2(c) there exists at least a value of η such that $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 \leq \eta$ when $(\mathbf{P1})$ is feasible. To see the necessity part, suppose that there exists an $\eta \in [0, \eta_{\max}]$ such that $\sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2 < \eta$. Then according to the fact that $\sum_{k=1}^K \|\mathbf{w}_k(0)\|_2^2 > 0$ and the concavity of $y = \sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$, there must exist an intersection point between $y = \sum_{k=1}^K \|\mathbf{w}_k(\eta)\|_2^2$ and $y = \eta$. The existence of such intersection point suggests that $(\mathbf{P1})$ is feasible and thus is a contradiction. ■

APPENDIX D

PROOF OF PROPOSITION 4

To prove the duality, we separate \mathbf{w}_k into the DL power p_k^D and beamforming direction $\tilde{\mathbf{w}}_k$, i.e., $\mathbf{w}_k = \sqrt{p_k^D} \tilde{\mathbf{w}}_k$ where $\|\tilde{\mathbf{w}}_k\|_2 = 1$. Then one can write (P2) in (37) as

$$\min_{\{\|\mathbf{w}_k\|_2=1\}} \left\{ \min_{\{p_k^D \geq 0\}, \{p_\ell^U \geq 0\}} \sum_{k=1}^K p_k^D (\tilde{\mathbf{w}}_k^H \mathbf{B} \tilde{\mathbf{w}}_k) + \sum_{\ell=1}^L p_\ell^U b_\ell \right. \quad (\text{A.21a})$$

$$\text{s.t. } \frac{p_k^D |\mathbf{h}_k^H \tilde{\mathbf{w}}_k|^2 / \rho_k^D}{\sum_{i=1}^K p_i^D \tilde{\mathbf{w}}_i^H \tilde{\mathbf{H}}_k \tilde{\mathbf{w}}_i + \hat{\sigma}_i^2(\mathbf{p}^U)} \geq 1, \quad \forall k \in \mathcal{K}, \quad (\text{A.21b})$$

$$\frac{p_\ell^U |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{k=1}^K p_k^D \tilde{\mathbf{w}}_k^H \mathbf{\Lambda}(\mathbf{v}_\ell, \mathbf{R}_{\Phi_0}) \tilde{\mathbf{w}}_k + \sum_{j=1}^L p_j^U \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \tilde{\sigma}_z^2} \geq 1, \quad \forall \ell \in \mathcal{L}. \quad (\text{A.21c})$$

Since the inner problem of (A.21) is a linear programming satisfying the Slater's condition, it has a zero duality gap with its Lagrange dual problem, which can be shown as

$$\max_{\boldsymbol{\lambda}, \boldsymbol{\mu} \geq \mathbf{0}} \boldsymbol{\lambda}^T \boldsymbol{\sigma}^2 + \tilde{\sigma}_z^2 \boldsymbol{\mu}^T \mathbf{1} \quad (\text{A.22a})$$

$$\text{s.t. } \frac{\lambda_k |\tilde{\mathbf{w}}_k^H \mathbf{h}_k|^2 / \rho_k^D}{\tilde{\mathbf{w}}_k^H \mathbf{Q}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \tilde{\mathbf{w}}_k} \leq 1, \quad k \in \mathcal{K}, \quad (\text{A.22b})$$

$$\frac{\mu_\ell |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L \mu_j \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \sum_{k=1}^K \lambda_k |f_{\ell k}|^2 + b_\ell} \leq 1, \quad \ell \in \mathcal{L}, \quad (\text{A.22c})$$

where $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_K]^T$ and $\boldsymbol{\mu} = [\mu_1, \dots, \mu_L]^T$ are respectively the dual variables associated with constraints (A.21b) and (A.21c). Now consider the following problem that is obtained by changing the 'max' to 'min' and ' \leq ' to ' \geq ' in (A.22)

$$\min_{\boldsymbol{\lambda}, \boldsymbol{\mu} \geq \mathbf{0}} \boldsymbol{\lambda}^T \boldsymbol{\sigma}^2 + \tilde{\sigma}_z^2 \boldsymbol{\mu}^T \mathbf{1} \quad (\text{A.23a})$$

$$\text{s.t. } \frac{\lambda_k |\tilde{\mathbf{w}}_k^H \mathbf{h}_k|^2 / \rho_k^D}{\tilde{\mathbf{w}}_k^H \mathbf{Q}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \tilde{\mathbf{w}}_k} \geq 1, \quad k \in \mathcal{K}, \quad (\text{A.23b})$$

$$\frac{\mu_\ell |\mathbf{v}_\ell^H \mathbf{g}_\ell|^2 / \rho_\ell^U}{\sum_{j=1}^L \mu_j \mathbf{v}_\ell^H \tilde{\mathbf{G}}_j \mathbf{v}_\ell + \sum_{k=1}^K \lambda_k |f_{\ell k}|^2 + b_\ell} \geq 1, \quad \ell \in \mathcal{L}. \quad (\text{A.23c})$$

It is easy to verify that constraints (A.22b) and (A.22c) of problem (A.22) hold with equality at the optimum. Therefore, problem (A.23) is feasible and constraints (A.23b) and (A.23c) also hold with equality at the optimum. Based on this fact, it is not difficult to verify that problems (A.22) and (A.23) have the same set of KKT conditions and therefore the two problems achieve the same optimal objective value. ■

REFERENCES

- [1] C.-H. Lee, T.-H. Chang, and S.-C. Lin, "Transmit-receive beamforming optimization for full-duplex cloud radio access networks," in *Proc. IEEE GLOBECOM*, Washington DC, USA, Dec 4-8, 2016, pp. 1–6.
- [2] R.-A. Pitaval, O. Tirkkonen, R. Wichman, K. Pajukoski, E. Lahetkangas., and E. Tirola, "Full-duplex self-backhauling for small-cell 5G networks," *IEEE Wireless Commun.*, pp. 83–89, Oct. 2015.
- [3] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sept. 2014.
- [4] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3702–3713, July 2012.
- [5] —, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1541–1553, Sept. 2012.
- [6] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.
- [7] M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. K. Ramakrishnan, C. W. Rice, and N. K. Shankaranarayanan, "Design and characterization of a full-duplex multi-antenna system for WiFi networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1160–1177, Mar. 2014.
- [8] D. Bharadia and S. Katti, "Full-duplex MIMO radios," in *Proc. USENIX NSDI*, Seattle, 2014, pp. 359–372.
- [9] S. Goyal, P. Liu, S. Panwar, R. A. DiFazio, R. Yang, and E. Bala, "Full duplex cellular systems: Will doubling interference prevent doubling capacity?" *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 121–127, May 2015.
- [10] A. C. Cirik, R. Wang, Y. Hua, and M. Latva-aho, "Weighted sum-rate maximization for full-duplex MIMO interference channels," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 801–815, March 2015.
- [11] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038–4050, Aug. 2013.
- [12] —, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4896–4910, Sep. 2014.
- [13] J. Kim, W. Choi, and H. Park, "Beamforming for full-duplex multiuser MIMO systems," to appear in *IEEE Trans. Vehicular Tech.*
- [14] S. Li, R. D. Murch, and V. K. Lau, "Linear transceiver design for full-duplex multi-user MIMO system," in *Proc. IEEE ICC*, Sydney, Australia, June 10-14, 2014, pp. 4921–4926.
- [15] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 5992–6010, Dec. 2013.
- [16] A. C. Cirik, O. Taghizadeh, R. Mathar, and T. Ratnarajah, "Qos considerations for full duplex multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 5, no. 1, pp. 36–39, Feb. 2016.
- [17] Y. Sun, D. W. K. Ng, J. Zhu, and R. Schober, "Multi-objective optimization for robust power efficient and secure full-duplex wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5511–5526, Aug. 2016.
- [18] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, pp. 62–75, May 2010.
- [19] A. Wiesel, Y. C. Eldar, and S. S. (Shitz), "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

- [20] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," Chapter 18 in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed., CRC Press, Aug. 2001.
- [21] Y.-F. Liu, M. Hong, and Y.-H. Dai, "Max-min fairness linear transceiver design problem for a multi-user SIMO interference channel is polynomial time solvable," *IEEE Signal Process. Lett.*, vol. 20, no. 1, pp. 27–30, Jan. 2013.
- [22] F. Rashid-Farrokh, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1449, Oct. 1998.
- [23] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," vol. 53, no. 1, pp. 18–28, Jan. 2004.
- [24] D. Bharadia, E. McMillin, and S. Katti, "Full duplex radios," in *Proc. ACM SIGCOMM*, pp. 375–386, 2013.
- [25] B. Song, R. L. Cruz, and B. D. Rao, "Network duality for multiuser MIMO beamforming networks and applications," *IEEE Trans. Commun.*, vol. 55, no. 3, pp. 618–631, March 2007.
- [26] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, pp. 20–34, May 2010.
- [27] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," in *Proc. IEEE VTC-Fall*, Vancouver, British Columbia, Canada, Sept. 24-29, 2002, pp. 87–91.
- [28] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE Veh. Technol. Conf.*, May 1999, p. 851856.
- [29] K.-K. Wong, G. Zheng, and T.-S. Ng, "Convergence analysis of downlink mimo antenna system using second-order cone programming," in *Proc. IEEE Veh. Technol. Conf.*, Sept. 2005, p. 492496.
- [30] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Max-min fairness linear transceiver design for a multi-user MIMO interference channel," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2413–2423, May 2013.
- [31] Q. Shi, M. Razavayan, M. Hong, and Z.-Q. Luo, "SINR constrained beamforming for a MIMO multi-user downlink system," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2920–2933, June 2016.
- [32] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, June 2007.
- [33] S. Boyd and A. Mutapcic, "Subgradient methods," available at www.stanford.edu/class/ee392o/subgrad_method.pdf.
- [34] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sept. 1995.
- [35] E. Bjornson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1807–1820, March 2010.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.